

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: July 19, 2005, 03:00:42 ; Search time 31 Seconds

(without alignments)  
3236.398 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074

Sequence: 1 MEGKSLTFFSYGLQWCLYE.....DDHSGVMSLGAAGLEGLVS 1344

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945.5	13.4	1454	US-09-388-221B-10	Sequence 10, Appl
2	945.5	13.4	1473	US-09-388-221B-2	Sequence 2, Appl
3	938.5	13.3	1399	US-09-388-221B-4	Sequence 4, Appl
4	938.5	13.3	1424	US-09-388-221B-12	Sequence 12, Appl
5	938.5	13.3	1443	US-09-388-221B-6	Sequence 6, Appl
6	630.5	8.9	456	US-08-910-731-4	Sequence 4, Appl
7	630.5	8.9	456	US-08-910-731-2	Sequence 2, Appl
8	621	8.8	456	US-08-795-395-4	Sequence 4, Appl
9	621	8.8	456	US-08-795-395-2	Sequence 2, Appl
10	612	8.7	456	US-08-910-731-8	Sequence 8, Appl
11	607.5	8.6	461	US-08-910-731-6	Sequence 6, Appl
12	374.5	5.3	953	US-09-245-281-43	Sequence 43, Appl
13	374.5	5.3	953	US-09-207-359B-43	Sequence 43, Appl
14	374.5	5.3	953	US-09-340-620A-43	Sequence 43, Appl
15	374.5	5.3	953	US-09-865-364-43	Sequence 43, Appl
16	364	5.1	1013	US-10-014-269-3	Sequence 3, Appl
17	364	5.1	1040	US-10-014-269-2	Sequence 2, Appl
18	364	5.1	1040	US-09-949-016-9080	Sequence 9080, Ap
19	363	5.1	1040	US-10-014-269-4	Sequence 4, Appl
20	356.5	5.0	209	US-09-388-221B-19	Sequence 19, Appl
21	347	4.9	953	US-09-099-041A-8	Sequence 8, Appl
22	347	4.9	953	US-09-245-281-8	Sequence 8, Appl
23	347	4.9	953	US-09-207-359B-8	Sequence 8, Appl
24	347	4.9	953	US-09-340-620A-8	Sequence 8, Appl
25	347	4.9	953	US-09-865-364-8	Sequence 8, Appl
26	345	4.9	1007	US-10-014-269-34	Sequence 34, Appl
27	324.5	4.6	966	US-09-207-359B-47	Sequence 47, Appl

28	324.5	4.6	966	US-09-865-364-47	Sequence 47, Appl
29	306.5	4.3	830	US-09-064-199-19	Sequence 19, Appl
30	306.5	4.3	1106	US-09-064-199-17	Sequence 17, Appl
31	306.5	4.3	1130	US-08-519-547A-6	Sequence 6, Appl
32	306.5	4.3	1130	US-09-064-199-18	Sequence 18, Appl
33	306.5	4.3	1207	US-09-064-199-16	Sequence 16, Appl
34	213	3.0	483	US-09-904-615-154	Sequence 154, App
35	163	2.3	305	US-10-014-269-7	Sequence 7, Appl
36	160	2.3	1240	US-08-930-996A-4	Sequence 4, Appl
37	153.5	2.2	490	US-09-099-041A-26	Sequence 26, Appl
38	153.5	2.2	490	US-09-245-281-26	Sequence 26, Appl
39	153.5	2.2	490	US-09-207-359B-26	Sequence 26, Appl
40	153.5	2.2	490	US-09-340-620A-26	Sequence 26, Appl
41	153.5	2.2	490	US-09-865-364-26	Sequence 26, Appl
42	149	2.1	509	US-08-809-999D-17	Sequence 17, Appl
43	149	2.1	509	US-09-069-637-17	Sequence 17, Appl
44	149	2.1	509	US-09-322-360-17	Sequence 17, Appl
45	149	2.1	509	US-09-131-831B-17	GENERAL INFORMA

## ALIGNMENTS

```
RESULT 1
US-09-388-221B-10
; Sequence 10, Application US/09388221B
; Patent No. 6618750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6618750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-11 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1454
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct
US-09-388-221B-10

Query Match      13.4% Score 945.5; DB 4; Length 1454;
Beet Local Similarity 24.8%; Pred. No. 5.7e-83;
Matches 294; Conservative 162; Mismatches 339; Gaps 30;

QY 11 SYGLQWCLYE-LDKREPFQFKELKKKS-SRSTCSIPQFEIENANVCEALLHVEYGA 68
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 5 AMGRACLTYLEFKKEBELKEFOLLANKAKHSSSGGETPAOPEKTSGMENVASYLVNQ-YGE 63
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 69 SLAVATSISENNMLRTLSKARD-----
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 64 QRAMDLAHTWEQNGRLCAQADGAGHSFPYSPSEPHLGSQPTSTAVLMPWTHE 123
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 95 -----MKKISQAM-----
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 124 LPAGCTQSSRRVLRQLPDTSGRRWRREISALYLQALPSSPDHSPSQESNAPTSTAVL 183
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 103 -----BOEGA-----PAETFEDELSQAMQDEBATAETEO 134
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 184 GSWSPPOPSLAPRBOEAPGTQWPLDSTGIIYTEIRERERKSKGRPPMAAVGTPQ 243
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 135 GHGG-----DTWDYKSHVWT-KF-----AEEEDVRSFENT 164
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 244 AHTSIQPHHHWEPVSRESLSTWPKNEDNQGFTOLLQORPHRSQDDLVKRS----- 299
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 165 AADWP-----EKQTLAG-AFDSDRNGFRPRVTVLHGKSGIGKSLARIVLCWA 212
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 300 ---PDDYVENRGHLIEIRDFGGLDTPQ---EPRIYILQGAAGIGKSTLAROYKEAWG 352
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 213 OGGLYQGMFSYVFLPVPVEMQKKSSVTEFISEKWPDSQAPVTEIMSRPRLFIIDGF 272
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

```

Db      353 RQGLYGRFQHFVFFSCELAQSKVVSIAELIGDKGTATPAPIRQILSRPERLLFILDGV 412
Qy      273 DDLSGVANN-DTKCKDMAEKOPFTLLRSILRKLLPESFLIYVVRGVTEKLSEVVS 331
Db      413 DEPGVNLQEPSESLCLHWSQOPADALLGSLGKLTLPBASPLIATATTAIONLPSLEQ 472
Qy      332 PRYLVRIGISEGRIHLLERIGIGEHOKTOGLRAIMNRRELLDQCVAVGSLICVALQL 391
Db      473 ARWVAVLGFSSSRKEFYRYFTDERQAIAPAFVYKSKKELMALCLVWVWSLACTCIMO 532
Qy      392 QDVVGSVAAPNQTLLG--LHAAFVHQUTPRGVVRCINLEERVVLCRPMAVEGVN 449
Db      533 QMKREKLTLSKTTTLCIHYLAQALQAPLG-----POLRDLCSIAEGIMQ 581
Qy      450 RKSVDGDLDLVQVGESEIRALFFMNTL----LPDSHCEYYTFPHLSLDDFCALYYV 505
Db      562 KKTLPSPDLRKHGIDGAIISTFLKMGILQEHPIPLS---YSFHLICFQDFPAAMSYV 636
Qy      506 LEGLE--IEPALCPLYVEKTKRSMELKQAGFIHSLM---MKRFLGLVSDVRRPLEVL 560
Db      637 LEDEKGRGKHSNCIIDEKLTLEA-----YGIHGLFGASTTRFLGLLSDGEREMENI 689
Qy      561 LGCPVPLGVKQKLIHWVSLG--QQPNAITPBGDTLDAPHCLFETODKEFVRLALNSPOEV 618
Db      690 FHCRLSQG--RNLQWVPSLQLLQPH-----SLESIHCLYETENKFTLQVMAHFPEM 741
Qy      619 WLPINQNDLIASSFCLOHCPYLKIRIVDKGIFPRDSAEACPVVPLMMDKTLIEBQW 678
Db      742 GWCETDMELVCTFCIKFSRHVKQLQ--IEG--RQHRSTWSPMTVVLFRRVPTDAYW 797
Qy      679 EDFCSMLGTHPHLRQLDLGSILITERAMKTLCAKLHPTCKIOTLMFRNAOITPGVQHLM 738
Db      798 -----QILF 801
Qy      739 RIWMANRULSLNCGTHLKEEDVYMACEBALKHPKCLLESRLDCCGLTHACYIKISQIL 798
Db      802 SVLKATRYMLKEILDLSGNSLSHSAVSKLCTKLRRPRCLLETLR----- 843
Qy      799 TTSPLKSLSLAGNKVTDQGVWPLSDALRVSCALQKLIIDCGITATGCOSSIASALVSN 858
Db      844 -----LAG-----CGITAECDKDLAFGIRAN 864
Qy      859 RSLTHCLSNNSLGNVNLCSRMRLPHCSLQRLMLNQCHLDTAGCGFLALALMGNSWL 918
Db      865 QT-----L 867
Qy      919 THLSISNMPVEDNGYKLLCEVWRBPSCHLOLELYKCHLTAAACESLSCVYSRHHKSL 978
Db      868 TELDLSFVNLTDGAKHLCORLRQPSCKLQRLQIVSCGLTSDCCODLASVLSASPSLKL 927
Qy      979 DLTDLALDGGVVALCEGLKQKNSVLTGLGKACGLTSDCCPAL 1022
Db      928 DLQONNLDVGVRLCEGLRHPACKLIRLGLDQTTLSDEMRQEL 971

RESULT 2
US-09-388-221B-2
; Sequence 2, Application US/09388221B
; Patent No. 6618750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6618750el Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221B-2

```

```

Query Match      13.4%; Score 945.5; DB 4; Length 1473;
Best Local Similarity 24.8%; Pred. No. 5.8e-83;
Matches 294; Conservative 162; Mismatches 339; Indels 389; Gaps 30;

Qy      11 SYGLQWCLYE-LDKEPOTFERELIKKKS-SESTTCSIFQEIENANVECLALLHEYYGA 68
Db      5 AMGLIACYLEFLKKEELKEFOLLANRAHSSSGETPPAPREKTSQMEVASYLVAQ--YGE 63
Qy      69 SLAATSIISIFENNALLTSEKARD----- 94
Db      64 QRADWLALHTWEGQLNSLCAQAOEGAGHSBSPFSPSEPHLSPSQPTSTAVLMPEWHE 123
Qy      95 -----MKKISQAM----- 102
Db      124 LPACQTGSSERRVLRQLPDTSGRRWRREISASLYQALPSPDHSPSQESPNAFTSTAVL 183
Qy      103 -----EFGCA-----TAATERQELISQAMEQCATAELEBQ 134
Db      184 GSWGSPQPSLAPREOAPGTOWPLDETSGIYYEIREREREKSEKGRPPMAAVVGTTPQ 243
Qy      135 GHGG-----DTWDYKSHWT--KF-----AEEEDVRRSPENT 164
Db      244 AHTSLQPHHHPWESVRESLCTWPMKQEDPNQKFTQULLQRPHPRSQDPLVKS----- 299
Qy      165 AADWP-----EMQTLAG--AFPSDRWGFPRFTVVLHGKSGIKSALARRIVLCWA 212
Db      300 ---WPDVYEENRGLHILRDLFGPGLDQ---EPRVILQAGAGIKSTLARQVKEAMG 352
Qy      213 QGGIYQGMFVFPFLPVREMKRKSSVTEFISHEMDSQAPVTEINSRERLLFITDGF 272
Db      353 RQGLYGRFQHFVFFSCELAQSKVVSIAELIGDKGTATPAPIRQILSRPERLLFILDGV 412
Qy      273 DDLSGVANN-DTKCKDMAEKOPFTLLRSILRKLLPESFLIYVVRGVTEKLSEVVS 331
Db      413 DEPGVNLQEPSESLCLHWSQOPADALLGSLGKLTLPBASPLIATATTAIONLPSLEQ 472
Qy      332 PRYLVRIGISEGRIHLLERIGIGEHOKTOGLRAIMNRRELLDQCVAVGSLICVALQL 391
Db      473 ARWVAVLGFSSSRKEFYRYFTDERQAIAPAFVYKSKKELMALCLVWVWSLACTCIMO 532
Qy      392 QDVVGSVAAPNQTLLG--LHAAFVHQUTPRGVVRCINLEERVVLCRPMAVEGVN 449
Db      533 QMKREKLTLSKTTTLCIHYLAQALQAPLG-----POLRDLCSIAEGIMQ 581
Qy      450 RKSVDGDLDLVQVGESEIRALFFMNTL----LPDSHCEYYTFPHLSLDDFCALYYV 505
Db      562 KKTLPSPDLRKHGIDGAIISTFLKMGILQEHPIPLS---YSFHLICFQDFPAAMSYV 636
Qy      506 LEGLE--IEPALCPLYVEKTKRSMELKQAGFIHSLM---MKRFLGLVSDVRRPLEVL 560
Db      637 LEDEKGRGKHSNCIIDEKLTLEA-----YGIHGLFGASTTRFLGLLSDGEREMENI 689
Qy      561 LGCPVPLGVKQKLIHWVSLG--QQPNAITPBGDTLDAPHCLFETODKEFVRLALNSPOEV 618
Db      690 FHCRLSQG--RNLQWVPSLQLLQPH-----SLESIHCLYETENKFTLQVMAHFPEM 741
Qy      619 WLPINQNDLIASSFCLOHCPYLKIRIVDKGIFPRDSAEACPVVPLMMDKTLIEBQW 678
Db      742 GWCETDMELVCTFCIKFSRHVKQLQ--IEG--RQHRSTWSPMTVVLFRRVPTDAYW 797
Qy      679 EDFCSMLGTHPHLRQLDLGSILITERAMKTLCAKLHPTCKIOTLMFRNAOITPGVQHLM 738
Db      798 -----QILF 801
Qy      739 RIWMANRULSLNCGTHLKEEDVYMACEBALKHPKCLLESRLDCCGLTHACYIKISQIL 798
Db      802 SVLKATRYMLKEILDLSGNSLSHSAVSKLCTKLRRPRCLLETLR----- 843
Qy      868 TELDLSFVNLTDGAKHLCORLRQPSCKLQRLQIVSCGLTSDCCODLASVLSASPSLKL 927
Db      928 DLQONNLDVGVRLCEGLRHPACKLIRLGLDQTTLSDEMRQEL 971
Qy      979 DLTDLALDGGVVALCEGLKQKNSVLTGLGKACGLTSDCCPAL 1022
Db      928 DLQONNLDVGVRLCEGLRHPACKLIRLGLDQTTLSDEMRQEL 971
Qy      859 RSLTHCLSNNSLGNVNLCSRMRLPHCSLQRLMLNQCHLDTAGCGFLALALMGNSWL 918

```



```
Db 865 QT-----L 867
Qy 919 THLSLNMVVDNGVKLLCEVWRERSCHLODELVKCHLTAACCSLSCVISRSRHLSL 978
Db 868 TELDSLSEFVLTDAAGAKHLQORLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKEI 927
Qy 979 DLTDNALGDGVAALCEGLKOKNSVLTRLGLKACGLTSDCCCEAL 1022
Db 928 DLQNNLDDVGVRLCEGLRHPACKLIRLGLDQTTLSDEKQEL 971

RESULT 3
US-09-388-221B-4
; Sequence 4, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-IJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221B-4

Query Match 13.3%; Score 938.5; DB 4; Length 1399;
Best Local Similarity 24.9%; Pred. No. 2,66-82;
Matches 291; Conservative 159; Mismatches 331; Indels 389; Gaps 30;

Qy 11 SYGLQWCLYE-LDKEEFOYFKELKKKS-SESTTCSIPQFIENANVECLALLHEYYGA 68
Db 5 AMGRACLAEFLFKELKEFOLLANKAHSRSSGETPAQPEKTSGMVASYLVAAQ-YGE 63
Qy 69 SLAWATSISENNNLRTLSKARD----- 94
Db 64 QRAMDLALHTWQGLRSLCAQAGAGHSPSPSPHILGSPSPSTAVLMPWTHE 123
Qy 95 -----MKKISQAM----- 102
Db 124 LPAGCTQSSERRVLAQLPDTSGRKREISASLLYALPSSPDHSPSPSPENAPTSTAVL 183
Qy 103 -----EOEGA-----TAETEOEISQAMEQEGTAETEOQ 134
Db 184 GSWSPSPQSLAPRQEAAPGTWPLDENTSIGIYTIIRERERKSKGRPPMAAVVGTTPQ 243
Qy 135 GHGG-----DTWDYKSHVMT-KF-----AEBEDVRRSFENT 164
Db 244 AHTSLQPHHWPESVRESLCSWPMKNEDFNQFTQLLLQRPHPRQODPLVRS----- 299
Qy 165 AADWP-----EMQTLAG-AFSDRWGFRPRVTVLHGKSGISGALARRIVLCMA 212
Db 300 ---WPDYVENRKHILIEIRDLFGPLDQ---EPRIVILQAGAGIGSTLARQYKEMWG 352
Qy 213 OGGLQAGPSYVFLPVRMORAKKSSVTEFISREMPQSAVPTIMSRPRELFIIDGF 272
Db 413 DEPGVLTQEPSESLCLHWSQPPADALIGSLIGKITLDEASFLITARTTALQNLIPSEIQ 472
Qy 273 DDLGSLVNN-DTKLCKDMAEKOPFTLIRSLIRKVLLEBSFLITVRDVGTEKSEVVS 331
Db 473 ARWEVILFSSSSREKRYRYFTDERQAIRAFRLVKSNNKEIMALCLVWVWSVLACTCTLMQ 532
Qy 332 PRYLVRISGEQRTHLLERIGIGHOKTQGLRAIMNRRELLDQOVAVAGSLICVALQL 391
Db 392 QDVVGESEVAPFNQTLTG--LHAAFVHQLTPRGVVRRCINLEERVVLKRFGMAVEGVNN 449
Db 533 QMKREKLTLSKTTTTCIHLTAQALQAPLG-----POLMDLCSLAEGITQ 581
```

```
Qy 450 RKSVPDGDLLWVQIGSESELALFHMNLT-----LPDSHCEBYTTFHLSLQDFCALYYV 505
Db 582 KTIILFSPDDLKRGHGDGAIIITFLKMGILQHPILS-----YSTHLCFQEFPAANSYV 636
Qy 506 LEGLE--TEPALCPDYVEKTRSRMELKQAGHISLW---MKRFLFGLVSDVRPRPLEVL 560
Db 637 LEDEGRKHSNCCIIDEKTEIA-----YGHGLFGASITTRFLGLLSBGEHEMENI 689
Qy 561 LCGPVPLGVKQKLLHWSVLG--QOPNATPBGTLDAFHLCEFTQDKEFVRLALNSFOEV 618
Db 690 FHCRLSQG--RNLQWVPISQLLLQPH-----SLESIHCIYETRNKFTQLQVMAHPEEM 741
Qy 619 WLPINQNLDIASSRCLOHCYLAIRKINDVGIFFRDESAACVVPVLMMDKTLIEQW 678
Db 742 GWCVTDEHLLVCTCFIKFSRHVKLDL-IEG---RQHRSTWSPTMVVLFPWVPVTDAYW 797
Qy 679 EDFCSMLGTHPHLQDLGSSILTERAMKTLCAKLRPTCKIQTLMFRNAQITPGVQHLM 738
Db 798 -----QLLF 801
Qy 739 RIVANNRILSLNIGTHLKEEDVPMACEALKHPKCLIESRLDCCGILTHACYLKSQIL 798
Db 802 SLAKTRILKELDLSGNSLSHSAVKSICKTLRRBRCLLETIR----- 843
Qy 799 TTSPLKSLSLAGKRVTDQGMPLSDALRVSCALQKILLEDGCGITATGCSLASALVSN 858
Db 844 -----LAG-----CGILDEDCKDLAFGLFAN 864
Qy 859 RSLTHLCLSNNSLAGEVNLICRSMRLPHCSLQRLMLNQCHLDTPAGCGFLALALMGNSWL 918
Db 865 QT-----L 867
Qy 919 THLSLNMVVDNGVKLLCEVWRERSCHLODELVKCHLTAACCSLSCVISRSRHLSL 978
Db 868 TELDSLSEFVLTDAAGAKHLQORLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKEI 927
Qy 979 DLTDNALGDGVAALCEGLKOKNSVLTRLGLKACGLTSDCCCEAL 1022
Db 928 DLQNNLDDVGVRLCEGLRHPACKLIRLGLDQTTLSDEKQEL 971

RESULT 4
US-09-388-221B-12
; Sequence 12, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-IJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1424
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct
US-09-388-221B-12

Query Match 13.3%; Score 938.5; DB 4; Length 1424;
Best Local Similarity 24.9%; Pred. No. 2,76-82;
Matches 291; Conservative 159; Mismatches 331; Indels 389; Gaps 30;

Qy 11 SYGLQWCLYE-LDKEEFOYFKELKKKS-SESTTCSIPQFIENANVECLALLHEYYGA 68
Db 5 AMGRACLAEFLFKELKEFOLLANKAHSRSSGETPAQPEKTSGMVASYLVAAQ-YGE 63
Qy 69 SLAWATSISENNNLRTLSKARD----- 94
Db 64 QRAMDLALHTWQGLRSLCAQAGAGHSPSPSPHILGSPSPSTAVLMPWTHE 123
```

```
OY -----MKKISQAM----- 102
Db 124 LPACGSGSERVYLRLPDTSGRRKREISASLILYQALPSSPDHSPSOSBPNAPTSTAVL 183
OY -----EODEGA-----TAETEOEISQAMEOGATAEBOQ 134
Db 184 GSWGSPQPSLAPRQEGAPGTOWPLDETSGIYTEIREREREKSEKGRPPMAAVGTPTPQ 243
OY GHGG-----DTWDYKSHWT-KF-----AEEEDVRSFENT 164
Db 244 AHTSLQPHHNPWBSVRESLCSYTWPMKXEDFNQKFTQLLLQRPHPRODPLVRS----- 299
OY AADWP-----EMQTLAG--AFPSDRWGFRPRTVVLHGSGIGSKALSARRIVLCMA 212
Db 300 ---WPDYVENRGHILIEIRDLFGSLDTQ---EPRIYILQGAAGIGSTILARQVKEAMG 352
OY QGGIYQGMFSYVFLPVREMOQKKESSVTEFISREWPDSQAPVTEIMSPERLLFITIDGF 272
Db 353 RGQLYGDRFQHFYFSCRELAQSKVSLAELIGKDGATAPAPIRQILISRPERLLFILDGV 412
OY DDLGSVLNN-DTKLCKMAKQRPPTLIRSLRKVLPESEFLIYTVRVGTREKLSSEVVS 331
Db 413 DEPGWVLOBPSESELCLHWSQOPADALGSLGKTLPEASFLITARTTALQNTLIPSELOQ 472
OY 332 PRYLLVRIGISEGORIHLLEBRIGEHOKTOGLRAIMNNREILDOCOVAVGSLICVALQ 391
Db 473 ARWVEVLGFSESSKREYFRYFTDERQALRAFLVYSKKEIMALCLVPMWSWLACTCLMQ 532
OY 392 QDVVGESVAPFNQTLTG--LHAAFFVHQLTTPRGVVRCLNLEERVVLKRFCEMAVEGVN 449
Db 533 QMKREKTLTTSKTTTTLCLHYLAQALQAPLG-----POLADLCSLAEGIMQ 581
OY 450 RKSVPDGDILMVQIGSESELRALFHMNTL---LPDSCEYTYTFPHLSIDPFCAALYV 505
Db 582 KKTLPSPDDLRKIGLDAIISTPLKMGILQENHPILS-----YSFHLCPQEFPAASVY 636
OY 506 LEGLE--IBPALCPLYVEKTKRSMELKQAGFHHSLSW---MKRFLFGLVSEDRVRPPEVL 560
Db 637 LEDEKGRGKHSNCTIIDLEKTLFA-----YGHGLFGASTTRFLGLISBEGREKEMNI 689
OY 561 LGCSPVLGVKOKLHMVSLLG--QOPNATTPGDTLDAFHLCTFETQDKEFVRLALNSFOEV 618
Db 690 FHCRLSQC--RNLQMOWPSLDLLOLQPH-----SLESILCLYETRNKFTLQVMAHFBEM 741
OY 619 WLPINQNDLILASSFCLOHCEYLRKIRVYKGIFFRDSASACPVPVPLMMDKTLIEBQW 678
Db 742 GMCVETDMLLVCTFCIKFSRHVKGLQI-IBG---ROHRSTWSPTMVVLFFRWVPTDLYW 797
OY 679 EDFCSMLGTHPHLRQDLGSSILTERAMKTCALKLHPTCKIQTLMFNNAQITPGVOHLM 738
Db 798 -----QLLF 801
OY 739 RIVANRNIRSLINSGTHLKEEDVYMACEALKNPKCLIESLRDCCGITHACYLKISQIL 798
Db 802 SVLKATRNILKEDELGSNLSHSAVKSILCKTLRRPRCLLETIR----- 843
OY 799 TTSFSLKSLSLAENKVTIQGVMPISDALRVSCALQKILIEDCGITATGCOSLASALVSN 858
Db 844 -----LAG-----CGITADCDKDLAFGLRAN 864
OY 859 RSLTHLCLSNNSLGNVNLICRSWRILPHSGLSQRLMLNQCHLDIADGCGFLALALMGNSWL 918
Db 865 QT-----L 867
OY 919 THLSLNNPVEDNGVKLCEVWRBPSCHLQDLIELVKCHLTAACESLSCVISRSRHLKSL 978
Db 868 TELDLSFVNLIDAGAKHLCQRLRQPSCKLQRLQIVSGCLTSDCCQDLASVLSASPSLKEI 927
OY 979 DLTNALADGGVAALCEGILKOKNSVLTIRLG 1008
Db 928 DLQONNLDVGVRLCEGLRHPACKLIRLG 957
```

```
RESULT 5
US-09-388-221B-6
; Sequence 6, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-1J 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221B
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221B-6

Query Match 13.3%; Score 938.5; DB 4; Length 1443;
Best Local Similarity 24.9%; Pred. No. 2,76-82;
Matches 291; Conservative 159; Mismatches 331; Indels 389; Gaps 30;

OY 11 SYGLQWCLYE-LDKKEPQTFKELKKKS-SESTTCSIPQFEIENAVNECLALLHEYYGA 68
Db 5 AMGRACYLEFLKEKEIKERFOLLANKAHSRSSGFTPADEKTSGMENVASYVAQ-YGE 63
OY 69 SLAATGISIFENNTLITSEKARD----- 94
Db 64 QRADWLALHTWQMGILSLCAQAOEGAGHSPSPSPSEPHLGSPOPTSTAVLMPWIEH 123
OY 95 -----MKKISQAM----- 102
Db 124 LPACGSGSERVYLRLPDTSGRRKREISASLILYQALPSSPDHSPSOSBPNAPTSTAVL 183
OY 103 -----EODEGA-----TAETEOEISQAMEOGATAEBOQ 134
Db 184 GSWGSPQPSLAPRQEGAPGTOWPLDETSGIYTEIREREREKSEKGRPPMAAVGTPTPQ 243
OY 135 GHGG-----DTWDYKSHWT-KF-----AEEEDVRSFENT 164
Db 244 AHTSLQPHHNPWBSVRESLCSYTWPMKXEDFNQKFTQLLLQRPHPRODPLVRS----- 299
OY 272 DDLGSVLNN-DTKLCKMAKQRPPTLIRSLRKVLPESEFLIYTVRVGTREKLSSEVVS 331
Db 413 DEPGWVLOBPSESELCLHWSQOPADALGSLGKTLPEASFLITARTTALQNTLIPSELOQ 472
OY 392 PRYLLVRIGISEGORIHLLEBRIGEHOKTOGLRAIMNNREILDOCOVAVGSLICVALQ 391
Db 473 ARWVEVLGFSESSKREYFRYFTDERQALRAFLVYSKKEIMALCLVPMWSWLACTCLMQ 532
OY 332 QDVVGESVAPFNQTLTG--LHAAFFVHQLTTPRGVVRCLNLEERVVLKRFCEMAVEGVN 449
Db 533 QMKREKTLTTSKTTTTLCLHYLAQALQAPLG-----POLADLCSLAEGIMQ 581
OY 450 RKSVPDGDILMVQIGSESELRALFHMNTL---LPDSCEYTYTFPHLSIDPFCAALYV 505
Db 582 KKTLPSPDDLRKIGLDAIISTPLKMGILQENHPILS-----YSFHLCPQEFPAASVY 636
OY 506 LEGLE--IBPALCPLYVEKTKRSMELKQAGFHHSLSW---MKRFLFGLVSEDRVRPPEVL 560
Db 637 LEDEKGRGKHSNCTIIDLEKTLFA-----YGHGLFGASTTRFLGLISBEGREKEMNI 689
OY 561 LGCSPVLGVKOKLHMVSLLG--QOPNATTPGDTLDAFHLCTFETQDKEFVRLALNSFOEV 618
Db 690 FHCRLSQC--RNLQMOWPSLDLLOLQPH-----SLESILCLYETRNKFTLQVMAHFBEM 741
OY 619 WLPINQNDLILASSFCLOHCEYLRKIRVYKGIFFRDSASACPVPVPLMMDKTLIEBQW 678
Db 742 GMCVETDMLLVCTFCIKFSRHVKGLQI-IBG---ROHRSTWSPTMVVLFFRWVPTDLYW 797
OY 679 EDFCSMLGTHPHLRQDLGSSILTERAMKTCALKLHPTCKIQTLMFNNAQITPGVOHLM 738
Db 798 -----QLLF 801
OY 739 RIVANRNIRSLINSGTHLKEEDVYMACEALKNPKCLIESLRDCCGITHACYLKISQIL 798
Db 802 SVLKATRNILKEDELGSNLSHSAVKSILCKTLRRPRCLLETIR----- 843
OY 799 TTSFSLKSLSLAENKVTIQGVMPISDALRVSCALQKILIEDCGITATGCOSLASALVSN 858
Db 844 -----LAG-----CGITADCDKDLAFGLRAN 864
OY 859 RSLTHLCLSNNSLGNVNLICRSWRILPHSGLSQRLMLNQCHLDIADGCGFLALALMGNSWL 918
Db 865 QT-----L 867
OY 919 THLSLNNPVEDNGVKLCEVWRBPSCHLQDLIELVKCHLTAACESLSCVISRSRHLKSL 978
Db 868 TELDLSFVNLIDAGAKHLCQRLRQPSCKLQRLQIVSGCLTSDCCQDLASVLSASPSLKEI 927
OY 979 DLTNALADGGVAALCEGILKOKNSVLTIRLG 1008
Db 928 DLQONNLDVGVRLCEGLRHPACKLIRLG 957
```

```

Db 690 FHCRLSGC--RNIQMWPVSLQLLQPH-----SLESILCLYETRNKFTLVQMAHFEEM 741
Qy 619 WLPINQNDLILASSRCLOHCPYLKIRVDVKIIFRDESAECPPVPLMRKTLIEBQM 678
Db 742 GNCVETDELLVCTCIRKFSRIVKQLQ-IEG---RQHRSTVSPTMVVLFRVWPVTDAYW 797
Qy 679 EDFCMLGTHPHRLQDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQTTPGVQHLN 738
Db 798 -----QILF 801
Qy 739 RIVMANRILRSINLGTHLKEEDVMAECALPKPKCLLESRLDCCGLTHACYLISQIL 798
Db 802 SVLKATRYMLKELDLSGNSLSHSAVSLCKTLARPRCLLETLR----- 843
Qy 799 TTSPLKSLISLAGNKVTDOGVWPLSDALRVSCALQKILBPGGTTATGCGSLASALVSN 858
Db 844 -----LAG-----CGLTAECCDLAAGLRAN 864
Qy 859 RSLTHCLSNNSLNEGAVNLCSRNLPHCSLQRLMLNQCHLDTAGCGFLALAMGNSWL 918
Db 865 QT-----L 867
Qy 919 THLSLWNPVEDNGVYKLLCEVNRBPSCHLDLELYKCHLTAACCSLSCVYSRRLKSL 978
Db 868 TELDSLFWLTDAGANKLQRLRPSCKLQRLQVSCGLTSDCCODIASVLSASPSLKL 927
Qy 979 DLTNALGDCGVYALCEGLKOKNSVYLRG 1008
Db 928 DLQNNLDVGVYRLCEGLRHPACKLIRIG 957

```

RESULT 6  
US-08-910-731-4  
Sequence 4, Application US/08910731  
Patent No. 5932440  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,731  
FILING DATE: (Herewith)  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,395  
FILING DATE: 04-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/794,546  
FILING DATE: 03-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0942.3440003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-731-4

```

Query Match 8.9%; Score 630.5; DB 2; Length 456;  
Best Local Similarity 34.0%; Pred. No. 8,5e-53;  
Matches 147; Conservative 75; Mismatches 201; Indels 9; Gaps 3;

```

Qy 659 EACPVPVPM-----MRDKTLIEBQWEDFCFMDGTHPHRLQDLGSSILTERAMKTLCAKLR 714
Db 17 ELPLPIQYQVVRDLDDCGLTEVRCKDIRSALQANPALTELSIRTELDAGVGLVQLQ 76
Qy 715 HPTCKIQTLMFRNAQIT-PGVQHLRIYMANRNLRSINLGTHLKEEDVMAECALKHPK 773
Db 77 NPTCKIQKLSLQNSCLTEAGCGVLPDVLRSLSTLRHLNDNPLDDEGLKILCEGLRBPQ 136
Qy 774 CLTESLRDCCGLTHACTYKLSQILTTSPSLKSLIAKNTVDGVMPFLDALRVSCAL 833
Db 137 CLTELQLEVCNLTATSCPEPLASVLRVPRDERELVLSNNDHEAGIHTLCOGLKDSACQL 196
Qy 834 OKLILEDGIGATYGOSLASALVSNRSLTHLCLSNNSLNEGAVNLCSRNLPHCSLQRL 893
Db 197 ESKLHENGITISANCKLDCVVASKASIQEIDLGSNKLGNIGIALCAGLLPSCKRLTL 256
Qy 894 MLNQCGLDTAGCGFLALAMGNSWLTHLSLWNPVEDNGVYKLLCEVNRBPSCHLDLELY 953
Db 257 WLMDCDVYABGCKDLCRYLRAKQSLKELSLAGNELKDEGAQLCEGLLEPGQLESIMWK 316
Qy 954 KCHLTAACCSLSCVYSRRLKSLDLTNALGDCGVYALCEGLKOKNSVYLRG 1013
Db 317 TCSLTAACCPHRCSTLYTRNSSLFELQMSNPNLGDGVVELCKALGYPTVLRVLMGDCD 376
Qy 1014 LTPDCEALSLALSCNRLTSLNLYONNFPSPKMKLCSAPACPSNTQIIGL-----WKM 1069
Db 377 VTDSGCSLATVLANRSLRELDLSNMGDNGVYQLLESIKQPSCLIQVLVLYDTWTMD 436
Qy 1070 QYPVOIRKLEE 1081
Db 437 EVEDQLRLEE 448

```

RESULT 7  
US-08-795-395-4  
Sequence 4, Application US/08795395  
Patent No. 5965399  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
TITLE OF INVENTION: Cloning and Expression of Rat Liver and  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,395  
FILING DATE: 04-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32, 893  
REFERENCE/DOCKET NUMBER: 0942.3440002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-795-395-4

Query Match 8.9%; Score 630.5; DB 2; Length 456;  
Best Local Similarity 34.0%; Pred. No. 8.5e-53;  
Matches 147; Conservative 75; Mismatches 201; Indels 9; Gaps 3;  
QY 659 EACPVPPLM---MRDKTLEEWEDFCSMGTGPHRLQDLGSSILTERAMKTLCAKLR 714  
DB 17 ELPLIQYVVRDLDGCLTEVRCKDIRSAIQANPALTELSTRTNELGDAVGIVLQGLQ 76  
QY 715 HPCTKIQTLMFRNAQIT-PGVQHLMRIYMANRNLSNLGGTHLKEEDVMACEALKHPK 773  
DB 77 NPTCKIQKLSIQNSLITFAGCGVLPDVLRSLSTRLRLIANDNPLGDEGLKTLCEGLRDPQ 136  
QY 774 CLLSRLDCCGGLTHACVLKISQILTTSPSLKSLSLGNAKKTLDGQVWPLSDALRVSOAL 833  
DB 137 CRLEKQLQLEYCNLTATSCBPALVLRKVPDFKEVLVSNNDHEAGIHITLCOGLKDSACQL 196  
QY 834 QKLLIEDCGITATGCGSLASALVSNRSLTHLCLSNNSLGNVALLCRSKMLPHCSIQRL 893  
DB 197 ESLKLENGCITSANCKDLDVVAASKASQELDLSNKLKNTGIALCSGILLPSCRLRTL 256  
QY 894 MLNQCILDTCAGCGPLALALMGNSWLTSLISMPVEDNGVRLCEVWREPSCHIQDELEV 953  
DB 257 WLMQCDVYAECKDLKRLVLRKQSLKELSLAGNELKDEGAQLLCESLIEPCQLESIMVK 316  
QY 954 KCHLTAACCELSLVISNRHLKSLDLTNALDGGVAALECGLKQKNSVYTRGLAKCG 1013  
DB 317 TCSLTAASCPHFCSVLTNSSLFELQWSSNPLGDSGVVLCALGYPTVLRVLMGDCD 376  
QY 1014 LTPSCCEALSLALSCNRHLTSLNLYVONFSPKMMKLCSAFACPTSNQIIGL---WKW 1069  
DB 377 VTDSGCSLAVLNLNRSLRELIDLSNMGMDNGVQLLESKQPSCTIQGLVLYDIWTYD 436  
QY 1070 QYVQVIRKLEE 1081  
DB 437 EVEDQLRALREE 448

RESULT 8  
US-08-910-731-2  
Sequence 2, Application US/08910731  
Patent No. 5932440  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
APPLICANT: SHANDILYA, HARINI  
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,731  
FILING DATE: (Herewith)  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,395  
FILING DATE: 04-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/794,546  
FILING DATE: 03-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32, 893  
REFERENCE/DOCKET NUMBER: 0942.3440003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-910-731-2

Query Match 8.8%; Score 621; DB 2; Length 456;  
Best Local Similarity 31.9%; Pred. No. 7.4e-52;  
Matches 151; Conservative 86; Mismatches 210; Indels 26; Gaps 5;  
QY 625 NLDIASFCLQHPYLRKIRVVDKGFPRDESAACPVPLM---MRDKTLEEWED 680  
DB 2 NLDI-----HCEQSDAR-----WTELLPILQYEVVRDLDGCLTEHCKD 42  
QY 681 FCSMLGTHPHRLQDLGSSILTERAMKTLCAKLRHPTCKIQITLMFRNAQIT-PGVQHLMR 739  
DB 43 IGSALRANPSSLTELCLATNELGDAVHLVQGLQSPCKIQKLSIQNSLITFAGCGVLP 102  
QY 740 IWMANRLSLNLGGTHLKEEDVMACEALKHPKCLLESIRLDCCGTHACYLKISQILT 799  
DB 103 TLRLSLPTLRHLSDNPLDAGRLCLCEGLLDPCHEKQLQLEVCRLTAASCEPLASVLR 162  
QY 800 TSPSLKSLSLGNAKKTLDGQVWPLSDALRVSOALQKLLIEDCGITATGCGSLASALVSNR 859  
DB 163 ATRALKELTVSNNDIGAGARVLAGGLADSAQQLETRLKNCGLTPANCDCGIVASQA 222  
QY 860 SLTHLCLSNNSLGNVALLCRSKMLPHCSIQRLMLNQCILDTCAGCGPLALALMGNSWLT 919  
DB 223 SLRELIDGNSGLDGAIAELCPGLLSPASRLKTLMLWECIDTASGCRDLKRLVQAKETLK 282  
QY 920 HLISMPVEDNGVRLCEVWREPSCHIQDELEVKCHLTAACCELSLVISNRHLKSLD 979  
DB 283 ELSLAGKRLDDEGARLLCESLLOPGCQLESIMVKSCSLTAACQOHSVLMITONHLELQ 342  
QY 980 LITNALDGGVAALECGLKQKNSVYTRGLKACGLTSDCCBALSLSCNRHLTSLNLYVQ 1039  
DB 343 LSSNKLDSGLOELCQALSQPGTTLRVLCIDGDEVYNSGSSILSLANRSLRELIDLSN 402  
QY 1040 NNFSPKMMKLCSAFACPTSNQIIGLMKWQYVQVIRKLEEYQLKP---RVV 1090  
DB 403 NCVGDDPVLQLLGLSLBQPCALBQVLYDYTMTEVEVEDRLQALEGSRGLRVI 455

RESULT 9  
US-08-795-395-2  
Sequence 2, Application US/08795395  
Patent No. 5965399  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
APPLICANT: SHANDILYA, HARINI  
TITLE OF INVENTION: Cloning and Expression of Rat Liver and

TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
 CITY: WASHINGTON  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3934  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/795,395  
 FILING DATE: 04-FEB-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/024,057  
 FILING DATE: 16-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ESMOND, ROBERT W.  
 REGISTRATION NUMBER: 32, 893  
 REFERENCE/DOCKET NUMBER: 0942.3440002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 456 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-795-395-2

Query Match 8.8%; Score 621; DB 2; Length 456;  
 Best Local Similarity 31.9%; Pred. No. 7.4e-52;  
 Matches 151; Conservative 86; Mismatches 210; Indels 26; Gaps 5;

625 NLDLIASSFCLOHCPYLRKIRVYVKGIFPRDSABACPVPLM---MRDKTLIEBQWED 680  
 2 NLDI-----HCQLSDAR-----WTELLPLDQYEVVRIDDDGLTEBHCKD 42

681 FCSMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIOTLMFRNAQIT-PCVQHLMR 739  
 43 IGSALRANPSTLTELCLRTNELGDAGVHLVLOGLSPTCKIQKLSLONCSLTEAGCGVLP 102

740 IVMANRNLRSINLGGTHLKEEDVMAACEALKHPRCLLSLRDCCGLTHACVLYKSQILT 799  
 103 TIRSLPTREHLSDNPLGADGRLRLCEGLDPCCHLEKQLECYRLTPAASCEPLASVLR 162

800 TSPSLKSLSLAGNKVTQGVWPLSDALRVSGCALQKILIEDCGITATGCGSLASALVSNR 859  
 163 ATRALKETLVSNNDIGEGARVILGQGLADSAQQLTILNENGLTPANCKDLGIVASQA 222

860 SLTHCLSNNSLIGNEGVNLCSRWRLPKCSLQRLMLNQCCHLDTAGCGFLALALMGNWLT 919  
 223 SLRELDLGSNGIGDAGIAELCPGLILSPASRLKTLMLMECDITPASCGRDLCRVLAQKETLK 282

920 HSLSMNIVENDGVLLCEVWREPSCHQDLELVYCHLTAACCELSLCVYSRHLKSLD 979  
 283 ELSLAKGDEGALLCESLQPCQLESIMVKSCLTPAACQHVSLMTQNKRLLELQ 342

980 LTDMALGQGVAAALCEGLKOKNSVLTGLKAKGTLSDCCALSLATLSNRLTSLNVQ 1039  
 343 LSSNLTGDSGIGELQALSQPETTLRVLCIDGCEVTNSGGCSLALILANSLSRLDLSN 402

1040 NNFPSKGMKLCSPAACPSTSNQIIGLMKWOYFVQIRKLLEEVQLKP-RVY 1090  
 403 NCVGDPGVLTQLGSLDEPGCALBOLVLYVTWTEVEEDRLQALBGSKGLRVT 455

RESULT 10  
 US-08-910-731-8  
 ; Sequence 8, Application US/08910731  
 ; Patent No. 5932440  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHATTERJEE, DEB K.  
 ; APPLICANT: SHANDILYA, HARINI  
 ; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
 ; CITY: WASHINGTON  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20005-3934  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/910,731  
 ; FILING DATE: (Herewith)  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/795,395  
 ; FILING DATE: 04-FEB-1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/794,546  
 ; FILING DATE: 03-FEB-1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/024,057  
 ; FILING DATE: 16-AUG-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: ESMOND, ROBERT W.  
 ; REGISTRATION NUMBER: 32, 893  
 ; REFERENCE/DOCKET NUMBER: 0942.3440003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 456 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: both  
 ; TOPOLOGY: both  
 ; MOLECULE TYPE: peptide  
 ; US-08-910-731-8

Query Match 8.7%; Score 612; DB 2; Length 456;  
 Best Local Similarity 32.1%; Pred. No. 5.7e-51;  
 Matches 152; Conservative 81; Mismatches 214; Indels 26; Gaps 5;

625 NLDLIASSFCLOHCPYLRKIRVYVKGIFPRDSABACPVPLM---MRDKTLIEBQWED 680  
 2 NLDI-----HCQLSDAR-----WTELLPLDQYEVVRIDDDGLTEBHCKD 42

681 FCSMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIOTLMFRNAQIT-PCVQHLMR 739  
 43 IGSALRANPSTLTELCLRTNELGDAGVHLVLOGLSPTCKIQKLSLONCSLTEAGCGVLP 102

740 IVMANRNLRSINLGGTHLKEEDVMAACEALKHPRCLLSLRDCCGLTHACVLYKSQILT 799  
 103 TIRSLPTREHLSDNPLGADGRLRLCEGLDPCCHLEKQLECYRLTPAASCEPLASVLR 162

800 TSPSLKSLSLAGNKVTQGVWPLSDALRVSGCALQKILIEDCGITATGCGSLASALVSNR 859  
 163 ATRALKETLVSNNDIGEGARVILGQGLADSAQQLTILNENGLTPANCKDLGIVASQA 222

860 SLTHCLSNNSLIGNEGVNLCSRWRLPKCSLQRLMLNQCCHLDTAGCGFLALALMGNWLT 919  
 223 SLRELDLGSNGIGDAGIAELCPGLILSPASRLKTLMLMECDITPASCGRDLCRVLAQKETLK 282







Db 834 EALK-DHPSLQRLMNOCHDITAGCGFLATALMGNSMLTHLSMNPVE-----DNGVKLL 936  
Qy 884 RLPCHSLQRLMNOCHDITAGCGFLATALMGNSMLTHLSMNPVE-----DNGVKLL 936  
Db 893 RV-NQTLRHLMWLIQNRITAKGTALRALQKNTAITETICLNGMLIKPEEAKVFEENERKII 951  
Qy 937 C 937  
Db 952 C 952

RESULT 14  
US-09-340-620A-43

; Sequence 43, Application US/09340620A  
; Patent No. 6482933  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-124001  
; CURRENT APPLICATION NUMBER: US/09/340,620A  
; CURRENT FILING DATE: 1999-06-28  
; PRIOR APPLICATION NUMBER: US 09/245,281  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: US 09/207,359  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: US 09/099,041  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: US 09/019,942  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-340-620A-43

Query Match 5.3%; Score 374.5; DB 4; Length 953;  
Best Local Similarity 22.7%; Pred. No. 6,2e-27;  
Matches 232; Conservative 150; Mismatches 398; Indels 241; Gaps 44;

Qy 54 NVECLA--LLIHEYYGASLAWATSIISIFENNNILRTLSKARDMKIKISQAMEOGATPAE 111  
Db 36 NTCCLVDNLLENYGFSA-----EDAEIVACAPTKDKYRKILDIYOSKGEVSE 84  
Qy 112 ---TEEOISQA-----MEOGATPAE-----TEEOGH--GGDTWDY 143  
Db 85 FFLYVLOQLLEDAVYDLRLMISEIGFSPQLIRTKITVNTDVSRYTQQLRHQIGRD---- 140  
Qy 144 KSHWTKFAEEDVYRSFENTAADPEMQTAG--AFPSDRMGF----- 185  
Db 141 -SKFMCLYAKQEDL--LLEET-----YNDTLMGLVGFNNEMLSIGSLDCLLDISTGYLN 192  
Qy 186 -RPRTVLHGKSGIGKSLARRIVLCMAOGGICYOGMESYVFLLPYREMORKES---SVT 241  
Db 193 EHGFEVFGDAGVGYKSWLQRLQSLMASGRL--TSTAKFFHFRCRMSCFKESDMLSIQ 251  
Qy 242 EFISREW----PDSQAPTEITMSRPERLLFIIDCFDDGS--VLANDTKLCMAAEKOPR 295  
Db 252 DLKFNHCYPRQDBEEVFSFLIRPHTALFTFDGIDELHSDPDISRVPDSCCPW--BPANP 310  
Qy 296 FTLLRSLLRKVLIPESFLIVTVR--DVGTEKLSEVVSPRYLLVYRGISGQRHL---- 348  
Db 311 LVLLANLISGRLLKAGKGLTARIGVEVPRQLRKV-----LIRGSPS--HLRAYA 361  
Qy 349 ---LLERIGIGHOKTOGLRAIMNNRELLDOCVAVGSLICVALQ--LDVVGESVAPENQ 404  
Db 362 RRMPEPRAQEHLL---LQGDANPNLCSLCGVPLFCMIIRCOHFGTVFEGSSSQLPD 417  
Qy 405 TLTGHLAAV-----FQGLTPRGVYRRCCLNEERVV-----LKRFRMAVVEGYMNRKS 452  
Db 418 CAVTLTDLVFLVTEVHLNRPOPSSILVORNTSPAEITLRAGWRTLHALGEVAHRTGDKSLF 477

Qy 453 VFDDDDLMVOGLGSESEIRLAFNNILPDSHCE--EYTFPHLSLODFCAALYYVLEGL 510  
Db 478 VFGGEVYQASKDLQEGDQLQFLR--ALPDVGPRGQSYEFPHLTLQAFPAFLVADDKV 535  
Qy 511 IEPALCPLYEYKTRSMELKQAGPHI-----HSLMKRPLFG 547  
Db 536 STRELIRFPREWTSPG--EATSSCHSFFSPQCLGGRSRLGPPFRNKDHFQFTNLFCVG 594  
Qy 548 LVSEDEVRRPLEVLLGCPVPGV---KOKLHWV-----SLGQOPNATTPG----- 590  
Db 595 LIKARQKILROL-----VPKAIIRRRKAL--WHLFASLSYLSLPRVOSGFGNOVHAM 649  
Qy 591 -DTLDAFHCLEFETQDKEFVRLALNSFOEVLPI---NONLDIASSFCLOHCPYLKIR 645  
Db 650 PTFLLMLRCIYETQSGVRLAARGISADYTLKAFACGADCSALSFVLHH--FHRQLA 707  
Qy 646 VDVGKIFPRDESACGVVLMRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERA 705  
Db 708 LIDLNNNINDYGVQ--ELQPCFSR-----LTVIRLSVNOQITDTG 744  
Qy 706 MKTLCARLHPTCKIOTLM--FRNAQITP--GVOHLMRIVMANRLIRSLNLGTHLKEEDVR 763  
Db 745 VKULCEEL--TKKIYTFGLYNNQITDGAARYAQIUDRCRGLKHLKLRKARITSEGGK 802  
Qy 764 MAGEALHKPKCLLESIRLDCCGLTHACYLKISQILTTPSPSLKSLAGNKVTDQGVNPLS 823  
Db 803 CVALAVN-----STSIVDVGMGNQIGDEGAKAFA 833  
Qy 824 DALRVSCALQKLILEDGITTATGCGSLASALVSNLSLTHLCISNNSLGEVNLICRSM 883  
Db 834 EALK-DHPSLQRLMNOCHDITAGCGFLATALMGNSMLTHLSMNPVE-----DNGVKLL 936  
Qy 884 RLPCHSLQRLMNOCHDITAGCGFLATALMGNSMLTHLSMNPVE-----DNGVKLL 936  
Db 893 RV-NQTLRHLMWLIQNRITAKGTALRALQKNTAITETICLNGMLIKPEEAKVFEENERKII 951  
Qy 937 C 937  
Db 952 C 952

RESULT 15  
US-09-865-364-43

; Sequence 43, Application US/09865364  
; Patent No. 6613521  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-112001  
; CURRENT APPLICATION NUMBER: US/09/865,364  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 09/207,359  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: US 09/099,041  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: US 09/019,942  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-865-364-43

Query Match 5.3%; Score 374.5; DB 4; Length 953;  
Best Local Similarity 22.7%; Pred. No. 6,2e-27;  
Matches 232; Conservative 150; Mismatches 398; Indels 241; Gaps 44;  
Qy 54 NVECLA--LLIHEYYGASLAWATSIISIFENNNILRTLSKARDMKIKISQAMEOGATPAE 111  
Db 36 NTCCLVDNLLENYGFSA-----EDAEIVACAPTKDKYRKILDIYOSKGEVSE 84

```
QY 112 ---TEORISOA-----MEOGATAAE-----TEEGH--GGDTWY 143
Db 85 PFLYVLOQLEDAVDRLMLSEIGSPSOLITKTIIVNTDPVSRYTQQLRHQGRD--- 140
QY 144 KSHVMTKPAEEDVRRSFENTAADPEMOTLAG--AFDSRWGF----- 185
Db 141 -SKFPLCYAQKEDL--LEET-----YMDTLMGLVGFNNENIGSLGDLCLDHSTGVLN 192
QY 186 -RPRTVLHGKSGIGKSALARIVLCWAOGGLYQGFVFLPVREWOKKES--SVT 241
Db 193 BHGETVVFVGDAVGKSMILQRLQSLMASGRL--TSTAKFFHFRCMFSCFESDMLSLQ 251
QY 242 BRISREW----PDSQAPTEIMSRPERLFIIDGPDLDGS--VLNNDTKLCKMAEKQRP 295
Db 252 DLFPHFCYPEODPEEVFSFLRPFHTALFTPDGLDELSDPDLRVPDSCCPW--EPAP 310
QY 296 FTLIRSLRKVLLPESFLIVTR---DVTEKLSKSEVSPRYLVRGISGEQRILH--- 348
Db 311 LVLLANLISGRLLKAGKLLTARTGVEVPROLLRKV-----LARGSPS---HLRAYA 361
QY 349 ---LIERGIGEHOKTOGLRAIMNREBLDQCOVPAGSLICVALQ-LQDVGESVAPFNQ 404
Db 362 RRMPEERTAOEHL-----LQQLDANPLCSLGGVPLFCWIIIFRCFQHFQTFEGSSSQLPD 417
QY 405 TLTLGHAAFV-----FHQLTGRGVRRCLNBERV-----LKRFCMAVEGYWNRKS 452
Db 418 CAVTLTDVFLVTEVHLNRPPSSLVQRNTRSPAETLRAGMRTLHALGEVNAHGTDKSLF 477
QY 453 VFDGDDLMVOGLGESELRALFMMNILLPDSHC--EYTFPHLSODPFCAALYYLEGLE 510
Db 478 VFGGEVQASKLOEGDLQGFUR--ALPDVGEQGSYEFPHLLOAFPTAFVLAADKV 535
QY 511 IEPALCPLYVEKTXSMELKQAGFI-----HSLMKRPLFG 547
Db 536 STRELLRFREWTSPG-EATSSCHSPSPFCGLGSRNLPDPFRNKDHQFTNLFPVG 594
QY 548 LVSEVRRRPLEVLLGCPPLGV--KQKILHV-----SLGQOPNATTPG----- 590
Db 595 LLAKAROKLLROL---VPKALIRRRKRL-WAHLFASLRSTKSLPRVOSGGFNQVHAM 649
QY 591 -DTLDAFCLFETODKEFVRLALNSFOEYWLPI---NONLDLIASSFLOHCPYLKIR 645
Db 650 PTELMMLACIYETQOKVGRLLAARGISADYLLAFCMACSDCSALFVLHH--FHROLA 707
QY 646 VDVKGI FPRDESAAACPVVPLMMRDKTLIEQWEDFCSMLGTHPHLROLGSSILTERA 705
Db 708 LDLDNNNNINDYGVQ--ELOPCFSR-----LTVIRLSVNQITDTG 744
QY 706 MKTILCAKLRHPTCKIQTLM-FRMAQITP-GVQHLMRYMANRLASLNLGTHLKEEDVR 763
Db 745 VKVLCEEL--TKYKLVTEFLGLYNNQITDIGARYAOIIDECRGLKHLKGNKRLTSEBGR 802
QY 764 MACBALKPKKLLBSLRLDCGGLTHACYLKISQIILTSPSLKSLSLAGNKVTDGVMPLS 823
Db 803 CVALLAVKN-----STSIYDVGMMGNQIGDBGAKAPA 833
QY 824 DALRVSOALQKLLIEDCGITATGCSLASALVSNRSLTHCLSNNSLGNBGNLCSRSM 883
Db 834 BALK-DHPSLITLTLAFLNGISPEGKSLAQLAKONTTLIVLTKNELNDBSABCFAMTL 892
QY 884 RLPHCSLQRLMNOCHLDTAGCGFLALALMGNMWLTHLSLNMVPE-----DNGVKLL 936
Db 893 RV-NQTLRHLMLQRIKTAKTALALOKNTAITEICLNGLIKPEBAKVFENEKRII 951
QY 937 C 937
Db 952 C 952
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 19, 2005, 11:19:57 ; Search time 475 Seconds  
(without alignments)  
1098.368 Million cell updates/sec

Title: US-10-066-521-6  
Perfect score: 7074  
Sequence: 1 MEGDKSLTFSSYGLQWCLVE.....DDHSGVMSLGAAGLEGLVS 1344

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1736639 seqs, 388188149 residues  
Total number of hits satisfying chosen parameters: 1736639

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7074	100.0	1344	US-10-124-498-6	Sequence 6, Appl1
2	7074	100.0	1344	US-10-066-521-6	Sequence 6, Appl1
3	5756	81.4	1162	US-10-216-645-2	Sequence 2, Appl1
4	5756	81.4	1162	US-10-416-642-1	Sequence 1, Appl1
5	5741.5	81.2	1200	US-10-399-443-24	Sequence 24, Appl1
6	5741.5	81.2	1200	US-10-677-943-24	Sequence 24, Appl1
7	5741.5	81.2	1200	US-10-860-761-4	Sequence 4, Appl1
8	5683.5	80.3	1143	US-10-116-645-4	Sequence 4, Appl1
9	5113.5	72.3	1033	US-10-092-900A-348	Sequence 348, Appl1
10	3318.5	46.9	682	US-10-407-866-96	Sequence 96, Appl1
11	2688	38.0	1111	US-10-216-645-5	Sequence 5, Appl1

	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	2577	2577	1940	1940	1885.5	1434	1434	1434	1434	1434	1402	1332.5	1332.5	1332.5	1320.5	1320.5	1314.5	1314.5	1313.5	1313.5	1303.5	1301.5	1289.5	1289.5	1283	1283	1255	1255	1255	1255	1255	1255	1255	1255
	36.4	36.4	27.4	27.4	26.7	20.3	20.3	20.3	20.3	20.3	19.8	18.8	18.8	18.7	18.7	18.7	18.6	18.6	18.6	18.6	18.4	18.4	18.2	18.2	18.1	18.1	17.7	17.7	17.7	17.7	17.7	17.7	17.7	17.6
	1111	1111	385	385	1093	994	994	994	994	994	994	1035	1035	1035	1035	1035	919	919	1036	1061	1061	1027	858	258	258	1032	1032	1034	1034	1034	1034	1034	1034	1062
	US-10-399-443-6	US-10-677-943-6	US-10-399-443-2	US-10-677-943-2	US-10-794-342-13	US-10-965-621-16	US-10-124-498-24	US-10-066-521-24	US-10-407-866-16	US-10-781-294-24	US-10-357-820-52	US-10-794-342-14	US-09-965-621-24	US-10-407-866-24	US-10-781-294-24	US-10-094-749-2718	US-10-161-493-4	US-10-398-037-7	US-10-794-342-18	US-10-124-498-18	US-10-066-521-18	US-10-407-866-68	US-09-848-035-8	US-09-986-224-8	US-10-399-443-4	US-10-677-943-4	US-10-882-761-33	US-10-127-516-5	US-10-027-629-5	US-10-028-374-18	US-10-132-967-5	US-10-183-770-18	US-10-264-9588-22	US-10-239-663-43
	Sequence 6, Appl1	Sequence 6, Appl1	Sequence 2, Appl1	Sequence 2, Appl1	Sequence 13, Appl1	Sequence 16, Appl1	Sequence 24, Appl1	Sequence 24, Appl1	Sequence 16, Appl1	Sequence 16, Appl1	Sequence 52, Appl1	Sequence 14, Appl1	Sequence 24, Appl1	Sequence 24, Appl1	Sequence 24, Appl1	Sequence 2718, Appl1	Sequence 8, Appl1	Sequence 7, Appl1	Sequence 18, Appl1	Sequence 18, Appl1	Sequence 18, Appl1	Sequence 8, Appl1	Sequence 8, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 3, Appl1	Sequence 3, Appl1	Sequence 5, Appl1	Sequence 5, Appl1	Sequence 18, Appl1	Sequence 18, Appl1	Sequence 22, Appl1	Sequence 43, Appl1	

## ALIGNMENTS

RESULT 1  
US-10-124-498-6  
Sequence 6, Application US/10124498  
Publication No. US20030017983A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Wang, Weiye  
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
FILE REFERENCE: 07334-367001  
CURRENT APPLICATION NUMBER: US/10/124,498  
PRIOR FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: 10/066,521  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/318,645  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 1344  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-124-498-6

Query Match 100.0%; Score 7074; DB 14; Length 1344;  
Beet Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 MEGDKSLTFSSYGLQWCLVEIDKEEFOTFKELKKSSBSTTCSIPQFIENAVVECAT 60  
Db 1 MEGDKSLTFSSYGLQWCLVEIDKEEFOTFKELKKSSBSTTCSIPQFIENAVVECAT 60

QY 61 LHHYYGASLAWATSIISIFENNNLRTLSEKARDPMKKSISQAMEOGATAETEEOEISQA 120  
DB 61 LHHYYGASLAWATSIISIFENNNLRTLSEKARDPMKKSISQAMEOGATAETEEOEISQA 120  
QY 121 MEQGATAETEEQGHGDTWYKSHVMTKFAEEDVRRSFENTADWPEMOTLAGAFDS 180  
DB 121 MEQGATAETEEQGHGDTWYKSHVMTKFAEEDVRRSFENTADWPEMOTLAGAFDS 180  
QY 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFPFVPEMOKKSSV 240  
DB 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFPFVPEMOKKSSV 240  
QY 241 TEFTSREWPDSQAPVTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPPTLIR 300  
DB 241 TEFTSREWPDSQAPVTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPPTLIR 300  
QY 301 SLRKVLLPESFLLIVYADVTEKLSKSEVSPRYLLVRGISGEORIHLLERGI GEHOKT 360  
DB 301 SLRKVLLPESFLLIVYADVTEKLSKSEVSPRYLLVRGISGEORIHLLERGI GEHOKT 360  
QY 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVGVESVAPFNQTLTGLHAAVFOHUTP 420  
DB 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVGVESVAPFNQTLTGLHAAVFOHUTP 420  
QY 421 RGVVRCCLNTEERVVLKRCFMAVEGVNRRKSVFDGDDLMVQGLGESELRALFFMNNILP 480  
DB 421 RGVVRCCLNTEERVVLKRCFMAVEGVNRRKSVFDGDDLMVQGLGESELRALFFMNNILP 480  
QY 481 DSHCEEYTFPHLSLODFCAALYYVLBGLTEIPALCPLYEKTYSMELKQAGFHISLW 540  
DB 481 DSHCEEYTFPHLSLODFCAALYYVLBGLTEIPALCPLYEKTYSMELKQAGFHISLW 540  
QY 541 MKRFLFGVSDVRRPLEVLLGCPVPLGVKOKLHMVSLGQONATPTGGTLDAFHCLF 600  
DB 541 MKRFLFGVSDVRRPLEVLLGCPVPLGVKOKLHMVSLGQONATPTGGTLDAFHCLF 600  
QY 601 ETODKEFRLALNFSFOEVMPLINQNLDIASSFCLOHCPYLRKIRVDYKGIFFPDESSEA 660  
DB 601 ETODKEFRLALNFSFOEVMPLINQNLDIASSFCLOHCPYLRKIRVDYKGIFFPDESSEA 660  
QY 661 CPVVPPLMRDXTLLIEBQWEDFCSMIGTAPHRLQDLGSSILTEBAMKTLCAKLHPCTCKI 720  
DB 661 CPVVPPLMRDXTLLIEBQWEDFCSMIGTAPHRLQDLGSSILTEBAMKTLCAKLHPCTCKI 720  
QY 721 QTLMEFRNAQITPGVOHLMRIYMANRNLRSLMGTHLKEEDVRRACEALKHPKCLBSLR 780  
DB 721 QTLMEFRNAQITPGVOHLMRIYMANRNLRSLMGTHLKEEDVRRACEALKHPKCLBSLR 780  
QY 781 LDCCGLTHACYLKTSQILITSPSLKSLSLAGNKYTDQVMPESDALRVSOQALQKLIED 840  
DB 781 LDCCGLTHACYLKTSQILITSPSLKSLSLAGNKYTDQVMPESDALRVSOQALQKLIED 840  
QY 841 CGITATGCGOSLASALVSNRSLTJHLCLSNNSLGNEGVNLGRSMRLPHGSLQRLMLNOCHL 900  
DB 841 CGITATGCGOSLASALVSNRSLTJHLCLSNNSLGNEGVNLGRSMRLPHGSLQRLMLNOCHL 900  
QY 901 DTACCGFLATLALMGNMWTJHLSLGMNPEVDNGVLLCEVWREBPSCHLODELIVYCHLTA 960  
DB 901 DTACCGFLATLALMGNMWTJHLSLGMNPEVDNGVLLCEVWREBPSCHLODELIVYCHLTA 960  
QY 961 CCESELSCVYSRSHKSLDLTDNALGDGVAAALCEGLKQKNSVLTJLGLKACGLTSDCE 1020  
DB 961 CCESELSCVYSRSHKSLDLTDNALGDGVAAALCEGLKQKNSVLTJLGLKACGLTSDCE 1020  
QY 1021 ALSLALSCNRHLTSLNIVONNFSPRKGMKCSAFACTSNLQIIGLWKMQYVVOIRKLE 1080  
DB 1021 ALSLALSCNRHLTSLNIVONNFSPRKGMKCSAFACTSNLQIIGLWKMQYVVOIRKLE 1080  
QY 1081 EVOLLKRRVVIDSGMHSFDEDDRHKIGLTFRLPESRAMPCALLMGNPEOKKRSLLAGD 1140  
DB 1081 EVOLLKRRVVIDSGMHSFDEDDRHKIGLTFRLPESRAMPCALLMGNPEOKKRSLLAGD 1140

QY 1141 FKSSTREFAKSLCIATANGESQVRDNNVQSSPOPAGTEHKKODKMLSVYSGAMSETAELE 1200  
DB 1141 FKSSTREFAKSLCIATANGESQVRDNNVQSSPOPAGTEHKKODKMLSVYSGAMSETAELE 1200  
QY 1201 GLGSGNSADHDHGGWAMSLGBLSRGICPTVLMTTAVCPGHMERLBSRGCLNSADHSG 1260  
DB 1201 GLGSGNSADHDHGGWAMSLGBLSRGICPTVLMTTAVCPGHMERLBSRGCLNSADHSG 1260  
QY 1261 VWSLGAAGLEGVLVSNADHSGVAMSLGAAGLEGVLVSNADHSGVAMSLGAAGLEGVL 1320  
DB 1261 VWSLGAAGLEGVLVSNADHSGVAMSLGAAGLEGVLVSNADHSGVAMSLGAAGLEGVL 1320  
QY 1321 SNSADHSGVAMSLGAAGLEGVLV 1344  
DB 1321 SNSADHSGVAMSLGAAGLEGVLV 1344

RESULT 2  
US-10-066-521-6  
: Sequence 6, Application US/1006521  
: Publication No. US20030027757A1  
: GENERAL INFORMATION:  
: APPLICANT: Bertin, John  
: APPLICANT: Wang, Weiye  
: APPLICANT: Blatcher, Maria  
: TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
: TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
: FILE REFERENCE: 07334-334001  
: CURRENT FILING DATE: 2002-06-25  
: PRIOR FILING DATE: 2001-09-10  
: PRIOR FILING DATE: 2001-09-10  
: PRIOR APPLICATION NUMBER: 60/265,231  
: NUMBER OF SEQ ID NOS: 25  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO 6  
: LENGTH: 1344  
: TYPE: PRN  
: ORGANISM: Homo sapiens  
US-10-066-521-6

Query Match 100.0%; Score 7074; DB 14; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGDKSLTFSSYGIOWCLYELDKBEFOTFKELLKKXSSESTTCSIPQFELNNVVECLAL 60  
DB 1 MEGDKSLTFSSYGIOWCLYELDKBEFOTFKELLKKXSSESTTCSIPQFELNNVVECLAL 60  
QY 61 LHHYYGASLAWATSIISIFENNNLRTLSEKARDPMKKSISQAMEOGATAETEEOEISQA 120  
DB 61 LHHYYGASLAWATSIISIFENNNLRTLSEKARDPMKKSISQAMEOGATAETEEOEISQA 120  
QY 121 MEQGATAETEEQGHGDTWYKSHVMTKFAEEDVRRSFENTADWPEMOTLAGAFDS 180  
DB 121 MEQGATAETEEQGHGDTWYKSHVMTKFAEEDVRRSFENTADWPEMOTLAGAFDS 180  
QY 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFPFVPEMOKKSSV 240  
DB 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFPFVPEMOKKSSV 240  
QY 241 TEFTSREWPDSQAPVTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPPTLIR 300  
DB 241 TEFTSREWPDSQAPVTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPPTLIR 300  
QY 301 SLRKVLLPESFLLIVYADVTEKLSKSEVSPRYLLVRGISGEORIHLLERGI GEHOKT 360  
DB 301 SLRKVLLPESFLLIVYADVTEKLSKSEVSPRYLLVRGISGEORIHLLERGI GEHOKT 360  
QY 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVGVESVAPFNQTLTGLHAAVFOHUTP 420  
DB 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVGVESVAPFNQTLTGLHAAVFOHUTP 420

```

QY 421 RGVARCINLEERYVLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLP 480
DB 421 RGVARCINLEERYVLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLP 480
QY 481 DSHCEYYTFFHLSDQFCALYYVLGELIEIPALCPLYVEKTKSMELKQAGFIHSLM 540
DB 481 DSHCEYYTFFHLSDQFCALYYVLGELIEIPALCPLYVEKTKSMELKQAGFIHSLM 540
QY 541 MKRFLFGLVSEVRRPLEVLLGCPVLYGKQKLLHWSVLLGQOPNATTPGDTLDAFHCLF 600
DB 541 MKRFLFGLVSEVRRPLEVLLGCPVLYGKQKLLHWSVLLGQOPNATTPGDTLDAFHCLF 600
QY 601 ETQDEPFRALANSFOEVLPIINQDLIASSFCLOHCPYLKIRVDVKGIFPRDESAA 660
DB 601 ETQDEPFRALANSFOEVLPIINQDLIASSFCLOHCPYLKIRVDVKGIFPRDESAA 660
QY 661 CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERAMKTLCAKLHPCTCKI 720
DB 661 CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERAMKTLCAKLHPCTCKI 720
QY 721 QTLMRNAQITPGVOHLWRIYMANRNLISLNGTHLKEEDVRMACEALKHPKCLBSLR 780
DB 721 QTLMRNAQITPGVOHLWRIYMANRNLISLNGTHLKEEDVRMACEALKHPKCLBSLR 780
QY 781 LDCCGLTHACVYKISQIILTTSPSLKSLAGNKVTDQGMPLSDALRVYSQCALQCLIED 840
DB 781 LDCCGLTHACVYKISQIILTTSPSLKSLAGNKVTDQGMPLSDALRVYSQCALQCLIED 840
QY 841 CGITATGCGSLASALVSNRSLTHLCLSNNSLGNENGLCRSMRLPHSLQRLMNOCHL 900
DB 841 CGITATGCGSLASALVSNRSLTHLCLSNNSLGNENGLCRSMRLPHSLQRLMNOCHL 900
QY 901 DTAGCGFLALAMGNSMTLTHLSLNNPVEDNGVKLLCEVMBRPSCHLQDLBLVKCHLTPA 960
DB 901 DTAGCGFLALAMGNSMTLTHLSLNNPVEDNGVKLLCEVMBRPSCHLQDLBLVKCHLTPA 960
QY 961 CCEBSLCTVSRSRHLKSLDLTNNALGDGVVALCEGLKQKNSVLTRLGLAKAGLTSDCCE 1020
DB 961 CCEBSLCTVSRSRHLKSLDLTNNALGDGVVALCEGLKQKNSVLTRLGLAKAGLTSDCCE 1020
QY 1021 ALSTLALSCNRHLTSLNLYONNFSPKMMKLSAFACPTSNLOIIGLMMKQVVOIRKLL 1080
DB 1021 ALSTLALSCNRHLTSLNLYONNFSPKMMKLSAFACPTSNLOIIGLMMKQVVOIRKLL 1080
QY 1081 EVOQLKPRVVIDGSWHSFDEDRHKIGLTFRLPESRAMP CALLMGNPEQKKRVSLLAGD 1140
DB 1081 EVOQLKPRVVIDGSWHSFDEDRHKIGLTFRLPESRAMP CALLMGNPEQKKRVSLLAGD 1140
QY 1141 FKSTRFAKSLCLATANGESQAVDNVEQSSPOPMAGTEHKODXMTSVGSGAMSETABLE 1200
DB 1141 FKSTRFAKSLCLATANGESQAVDNVEQSSPOPMAGTEHKODXMTSVGSGAMSETABLE 1200
QY 1201 GLGNSADHDGGMASLGRRLSRLGCTPYVMTTAVCPGHWERLGSRGWCLNSADDSHG 1260
DB 1201 GLGNSADHDGGMASLGRRLSRLGCTPYVMTTAVCPGHWERLGSRGWCLNSADDSHG 1260
QY 1261 VWSWLSAAGLEGLVNSADDSHGVMASLGAAGLEGLVNSADDSHGVMASLGAAGLEGLV 1320
DB 1261 VWSWLSAAGLEGLVNSADDSHGVMASLGAAGLEGLVNSADDSHGVMASLGAAGLEGLV 1320
QY 1321 SNSADDSHGVMASLGAAGLEGLV 1344
DB 1321 SNSADDSHGVMASLGAAGLEGLV 1344

```

```

RESULT 3
US-10-216-645-2
; Sequence 2, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA

```

```

; APPLICANT: PETERS-KOTTIG, MICHAEL
; APPLICANT: BECKMANN, GORG
; TITLE OF INVENTION: HUMAN MATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216, 645
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-645-2

Query Match      81.4%; Score 5756; DB 14; Length 1162;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1102; Conservative 2; Mismatches 0; Indels 54; Gaps 1;

QY 1 MEGDKSLTFSSYGLQWCLYEIDKEEFQTFKELKKKSESTTCSIPOFEINAVECAL 60
DB 1 MEGDKSLTFSSYGLQWCLYEIDKEEFQTFKELKKKSESTTCSIPOFEINAVECAL 60
QY 61 LHEHYGASLMAATSIIFENMNLRTLSEKARDMK----- 96
DB 61 LHEHYGASLMAATSIIFENMNLRTLSEKARDMKSPEDPATMTDQPSKEKEVPENK 120
QY 97 -----KISQMEQEGATAETEOEISQMEQEGA 126
DB 97 -----KISQMEQEGATAETEOEISQMEQEGA 126
QY 121 YGMTKLIGVSDISDNNKHKYVGHSFALISQMEQEGATAETEOEISQMEQEGA 180
DB 121 YGMTKLIGVSDISDNNKHKYVGHSFALISQMEQEGATAETEOEISQMEQEGA 180
QY 127 TAAETEOEGCGDWDYKSHVMTKFAEEDVRSFENTADWPENQTLAGAFSDRWGFR 186
DB 127 TAAETEOEGCGDWDYKSHVMTKFAEEDVRSFENTADWPENQTLAGAFSDRWGFR 186
QY 181 TAAETEOEGCGDWDYKSHVMTKFAEEDVRSFENTADWPENQTLAGAFSDRWGFR 240
DB 181 TAAETEOEGCGDWDYKSHVMTKFAEEDVRSFENTADWPENQTLAGAFSDRWGFR 240
QY 187 PRTVLHGKSGIGSALARRIVLCMAOGLYQGMPSYVFLPVEMORKKSSVTEFTSR 246
DB 187 PRTVLHGKSGIGSALARRIVLCMAOGLYQGMPSYVFLPVEMORKKSSVTEFTSR 246
QY 241 PRTVLHGKSGIGSALARRIVLCMAOGLYQGMPSYVFLPVEMORKKSSVTEFTSR 300
DB 241 PRTVLHGKSGIGSALARRIVLCMAOGLYQGMPSYVFLPVEMORKKSSVTEFTSR 300
QY 247 EMPDQAPVTEIMSRPERLFIIDFPDLGSLVNDTCLCDMAEKOPFTLIRSLRKV 306
DB 247 EMPDQAPVTEIMSRPERLFIIDFPDLGSLVNDTCLCDMAEKOPFTLIRSLRKV 306
QY 301 EMPDQAPVTEIMSRPERLFIIDFPDLGSLVNDTCLCDMAEKOPFTLIRSLRKV 360
DB 301 EMPDQAPVTEIMSRPERLFIIDFPDLGSLVNDTCLCDMAEKOPFTLIRSLRKV 360
QY 307 LLPSEFLIVTVRDVTEKLSKEVVSPLYLVIRGISGEORIHLLERIGEGHOKTQGLRAI 366
DB 307 LLPSEFLIVTVRDVTEKLSKEVVSPLYLVIRGISGEORIHLLERIGEGHOKTQGLRAI 366
QY 361 LLPSEFLIVTVRDVTEKLSKEVVSPLYLVIRGISGEORIHLLERIGEGHOKTQGLRAI 420
DB 361 LLPSEFLIVTVRDVTEKLSKEVVSPLYLVIRGISGEORIHLLERIGEGHOKTQGLRAI 420
QY 367 MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFFHQLTRGVRR 426
DB 367 MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFFHQLTRGVRR 426
QY 421 MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFFHQLTRGVRR 480
DB 421 MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFFHQLTRGVRR 480
QY 427 CLNLEERYVLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 486
DB 427 CLNLEERYVLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 486
QY 481 CLNLEERYVLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 540
DB 481 CLNLEERYVLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 540
QY 487 YTFEFLSLDQFCALYYVLGELIEIPALCPLYVEKTKSMELKQAGFIHSLMKEFL 546
DB 487 YTFEFLSLDQFCALYYVLGELIEIPALCPLYVEKTKSMELKQAGFIHSLMKEFL 546
QY 541 YTFEFLSLDQFCALYYVLGELIEIPALCPLYVEKTKSMELKQAGFIHSLMKEFL 600
DB 541 YTFEFLSLDQFCALYYVLGELIEIPALCPLYVEKTKSMELKQAGFIHSLMKEFL 600
QY 547 GLVSEVRRPLEVLLGCPVLYGKQKLLHWSVLLGQOPNATTPGDTLDAFHCLFETQDKE 606
DB 547 GLVSEVRRPLEVLLGCPVLYGKQKLLHWSVLLGQOPNATTPGDTLDAFHCLFETQDKE 606
QY 601 GLVSEVRRPLEVLLGCPVLYGKQKLLHWSVLLGQOPNATTPGDTLDAFHCLFETQDKE 660
DB 601 GLVSEVRRPLEVLLGCPVLYGKQKLLHWSVLLGQOPNATTPGDTLDAFHCLFETQDKE 660
QY 607 FVRLALNSFOEVLPIINQDLIASSFCLOHCPYLKIRVDVKGIFPRDESAAECPPVPL 666
DB 607 FVRLALNSFOEVLPIINQDLIASSFCLOHCPYLKIRVDVKGIFPRDESAAECPPVPL 666
QY 661 FVRLALNSFOEVLPIINQDLIASSFCLOHCPYLKIRVDVKGIFPRDESAAECPPVPL 720
DB 661 FVRLALNSFOEVLPIINQDLIASSFCLOHCPYLKIRVDVKGIFPRDESAAECPPVPL 720
QY 667 WMRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERAMKTLCAKLHPCTCKIQTLMFR 726
DB 667 WMRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERAMKTLCAKLHPCTCKIQTLMFR 726
QY 721 WMRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERAMKTLCAKLHPCTCKIQTLMFR 780
DB 721 WMRDKTLIEQWEDFCSMLGTHPHLRQLDGLSSILTERAMKTLCAKLHPCTCKIQTLMFR 780
QY 727 NAQITPGVOHLWRIYMANRNLISLNGTHLKEEDVRMACEALKHPKCLBSLRDCCGL 786
DB 727 NAQITPGVOHLWRIYMANRNLISLNGTHLKEEDVRMACEALKHPKCLBSLRDCCGL 786

```

Db 781 NAQITPGVQHLMRIYMANRNRSLNIGSTHLKEEDVMAACEALKHFKCLLESRLDCCGL 840  
Qy 787 THACVTKISQILTTSPSKSLISLGNKVTDOGWMPLSALPVSOCALOKLLEDCGTTAT 846  
Db 841 THACVTKISQILTTSPSKSLISLGNKVTDOGWMPLSALPVSOCALOKLLEDCGTTAT 900  
Qy 847 GCOSLASALVNSRSLTHLCLSNNSLGNENVALLCRSNRLPHCSLQRLMLNQCCHLDTAGCG 906  
Db 901 GCOSLASALVNSRSLTHLCLSNNSLGNENVALLCRSNRLPHCSLQRLMLNQCCHLDTAGCG 960  
Qy 907 FLALALMGNSWLTHTLSLGNPVEDNGVYKLCVWRBPSCHLODELVYKCHLTAACCSLS 966  
Db 961 FLALALMGNSWLTHTLSLGNPVEDNGVYKLCVWRBPSCHLODELVYKCHLTAACCSLS 1020  
Qy 967 CVISRSHLKSLLDTDNALGQGVAAALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1026  
Db 1021 CVISRSHLKSLLDTDNALGQGVAAALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1080  
Qy 1027 SCNRHLTSLNLYONNFPSPKGMKLCSAFACPTSNLQIIGLWKQYPVQIRKLEEVOLK 1086  
Db 1081 SCNRHLTSLNLYONNFPSPKGMKLCSAFACPTSNLQIIGLWKQYPVQIRKLEEVOLK 1140  
Qy 1087 PRVVIDGSMHSPDEDDRH 1104  
Db 1141 PRVVIDGSMHSPDEDDRY 1158

RESULT 4  
US-10-416-642-1  
; Sequence 1, Application US/10416642  
; Publication No. US20040043452A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: RAMKUMAR, Jayalaxmi  
; APPLICANT: ARVIZU, Chandra  
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0842 PCT  
; CURRENT APPLICATION NUMBER: US/10/416,642  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/249,407  
; PRIOR FILING DATE: 2000-11-15  
; NUMBER OF SEQ. ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 1162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CD1  
US-10-416-642-1

Query Match 81.4%; Score 5756; DB 15; Length 1162;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 1102; Conservative 2; Mismatches 0; Indels 54; Gaps 1;  
Qy 1 MEGKSLTFSSYGLQWCLYELDKERFOTFKELKKKSESTTCGIPQEIENAVECAL 60  
Db 1 MEGKSLTFSSYGLQWCLYELDKERFOTFKELKKKSESTTCGIPQEIENAVECAL 60  
Qy 61 LHHEYYGASLMAWTSISIFENNNLRTLSEKARDMK----- 96  
Db 61 LHHEYYGASLMAWTSISIFENNNLRTLSEKARDMK----- 96  
Qy 97 -----KISQAMEQEGATAAETEOISQAMEQGA 126  
Db 121 YGATKILIVGSDISDNNKHKYGVIHSSGFATISQAMEQEGATAAETEOISQAMEQGA 180  
Qy 127 TAAETEOGCHGDDWYKSHMTKFAEEEDVRSBPENTAAAMPQMTLAGAFDSDRNGFR 186  
Db 181 TAAETEOGCHGDDWYKSHMTKFAEEEDVRSBPENTAAAMPQMTLAGAFDSDRNGFR 240

Qy 187 PRVVLHGKSGIGKSALARRIVLCWAQGGIYQGMFSYVFFLPVREMRKKESSVTEFISR 246  
Db 241 PRVVLHGKSGIGKSALARRIVLCWAQGGIYQGMFSYVFFLPVREMRKKESSVTEFISR 300  
Qy 247 EWPDSQAPYTEINSRPERLLFIIDGPDLLGSVLANPDKLCKDNAEKPPFTLLRSLLRV 306  
Db 301 EWPDSQAPYTEINSRPERLLFIIDGPDLLGSVLANPDKLCKDNAEKPPFTLLRSLLRV 360  
Qy 307 LLPESFLIVVRDVGTEKLSEVYSPRYLLVRGISEGORIHLLEERIGIHEKOTQGLRAI 366  
Db 361 LLPESFLIVVRDVGTEKLSEVYSPRYLLVRGISEGORIHLLEERIGIHEKOTQGLRAI 420  
Qy 367 MNNRELLDQCVPAVGSILCVALQLODVESVAPFNQVLTGLAAAFVHQLTFRGVRR 426  
Db 421 MNNRELLDQCVPAVGSILCVALQLODVESVAPFNQVLTGLAAAFVHQLTFRGVRR 480  
Qy 427 CLNLEERVYTKRFRMAVEGVNMRKSPDDGDLVAVQLGSEELRALPHNNIILPDSICEE 486  
Db 481 CLNLEERVYTKRFRMAVEGVNMRKSPDDGDLVAVQLGSEELRALPHNNIILPDSICEE 540  
Qy 487 YTFPHLSLODFCALYYVLEGLIEIPALCPLYEKTFRSMELKQAGFTHSLMKRFLF 546  
Db 541 YTFPHLSLODFCALYYVLEGLIEIPALCPLYEKTFRSMELKQAGFTHSLMKRFLF 600  
Qy 547 GLVSEDEVRRPLEVLLAGCPVPLGVYKQLLHWVSLGQGPNAITPBDTLDAFHCLFETQDXE 606  
Db 601 GLVSEDEVRRPLEVLLAGCPVPLGVYKQLLHWVSLGQGPNAITPBDTLDAFHCLFETQDXE 660  
Qy 607 FVRALNSFOEWLPINQNDLILASSFCLOHCPYLKIRIVDVNGIFPRDSSAECPPVPL 666  
Db 661 FVRALNSFOEWLPINQNDLILASSFCLOHCPYLKIRIVDVNGIFPRDSSAECPPVPL 720  
Qy 667 WMRDKTLIEBOWEDFCSMLGTHPHLRQDLGSSILTEBAMKTYLCAKLRHTCKIQTLMFR 726  
Db 721 WMRDKTLIEBOWEDFCSMLGTHPHLRQDLGSSILTEBAMKTYLCAKLRHTCKIQTLMFR 780  
Qy 727 NAQITPGVQHLMRIYMANRNRSLNIGSTHLKEEDVMAACEALKHFKCLLESRLDCCGL 786  
Db 781 NAQITPGVQHLMRIYMANRNRSLNIGSTHLKEEDVMAACEALKHFKCLLESRLDCCGL 840  
Qy 787 THACVTKISQILTTSPSKSLISLGNKVTDOGWMPLSALPVSOCALOKLLEDCGTTAT 846  
Db 841 THACVTKISQILTTSPSKSLISLGNKVTDOGWMPLSALPVSOCALOKLLEDCGTTAT 900  
Qy 847 GCOSLASALVNSRSLTHLCLSNNSLGNENVALLCRSNRLPHCSLQRLMLNQCCHLDTAGCG 906  
Db 901 GCOSLASALVNSRSLTHLCLSNNSLGNENVALLCRSNRLPHCSLQRLMLNQCCHLDTAGCG 960  
Qy 907 FLALALMGNSWLTHTLSLGNPVEDNGVYKLCVWRBPSCHLODELVYKCHLTAACCSLS 966  
Db 961 FLALALMGNSWLTHTLSLGNPVEDNGVYKLCVWRBPSCHLODELVYKCHLTAACCSLS 1020  
Qy 967 CVISRSHLKSLLDTDNALGQGVAAALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1026  
Db 1021 CVISRSHLKSLLDTDNALGQGVAAALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1080  
Qy 1027 SCNRHLTSLNLYONNFPSPKGMKLCSAFACPTSNLQIIGLWKQYPVQIRKLEEVOLK 1086  
Db 1081 SCNRHLTSLNLYONNFPSPKGMKLCSAFACPTSNLQIIGLWKQYPVQIRKLEEVOLK 1140  
Qy 1087 PRVVIDGSMHSPDEDDRH 1104  
Db 1141 PRVVIDGSMHSPDEDDRY 1158

RESULT 5  
US-10-399-443-24  
; Sequence 24, Application US/10399443  
; Publication No. US20040028669A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; Secretary, Department of Health & Human Services, The National Institute c  
; APPLICANT: Health



; APPLICANT: Nelson, Lawrence M.  
 ; APPLICANT: Tong, Zhi-Bin  
 ; APPLICANT: Nelson, Lawrence  
 ; APPLICANT: Zhi-Bin, Tong  
 ; TITLE OF INVENTION: Human Gene Critical to Fertility  
 ; FILE REFERENCE: 4239-64785  
 ; CURRENT APPLICATION NUMBER: US/10/399,443  
 ; PRIOR FILING DATE: 2003-04-16  
 ; PRIOR APPLICATION NUMBER: 60/241,510  
 ; PRIOR FILING DATE: 2000-10-18  
 ; PRIOR APPLICATION NUMBER: PCT/US01/10981  
 ; PRIOR FILING DATE: 2001-04-04  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 24  
 ; LENGTH: 1200  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-399-443-24

Query Match 81.2%; Score 5741.5; DB 15; Length 1200;  
 Best Local Similarity 95.9%; Pred. No. 0;  
 Matches 1098; Conservative 3; Mismatches 3; Indels 41; Gaps 1;  
 1 MEGKSLTFSSYGLQWCLYELDKKEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 60  
 52 MEGKSLTFSSYGLQWCLYELDKKEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 111  
 61 LHHEYGASLAWATSIISIFENMNLRTLSEKARDMKK----- 97  
 112 LHHEYGASLAWATSIISIFENMNLRTLSEKARDMKKSHSPDEPATWTDGPKSEKVPGI 171  
 98 -----ISQAMEBEGATAETBEQSIQAMEBEGATAETBEQSGCD 139  
 172 SQAVQDSATAETKEQSIQAMEBEGATAETBEQSIQAMEBEGATAETBEQSGCD 231  
 140 TWDKSHMTKFAEBEDVRSFENTAAADMPMOTLAGAFSDRGMGFRRTVVLHGKSGIG 199  
 232 TWDKSHMTKFAEBEDVRSFENTAAADMPMOTLAGAFSDRGMGFRRTVVLHGKSGIG 291  
 200 KSALARIIVLCMAOGLYQGMFSYVFLPVREMOCKESSYTEFISREMPDSQAPVEIM 259  
 292 KSALARIIVLCMAOGLYQGMFSYVFLPVREMOCKESSYTEFISREMPDSQAPVEIM 351  
 260 SRPERLFIIDGFDLGSVLANDTYLCQMAKOPPLTSLRLKVLPPESFLIVYRD 319  
 352 SRPERLFIIDGFDLGSVLANDTYLCQMAKOPPLTSLRLKVLPPESFLIVYRD 411  
 320 VGTETKSEVSPRLIVRGISGEORIHLLERGI GEHOKTQGLRAIMNNEELDQCQVP 379  
 412 VGTETKSEVSPRLIVRGISGEORIHLLERGI GEHOKTQGLRAIMNNEELDQCQVP 471  
 380 AVGSILCYALQLODVGVESVAPFNQTLTGLHAAFFHOLTGRGVARRCLNTEERVYLRKF 439  
 472 AVGSILCYALQLODVGVESVAPFNQTLTGLHAAFFHOLTGRGVARRCLNTEERVYLRKF 531  
 440 CRMAVEGWNKRSYVFDGDLWVGIGSESELRALPHMNLILPDSHCEEYTFPHLSLDQFC 499  
 532 CRMAVEGWNKRSYVFDGDLWVGIGSESELRALPHMNLILPDSHCEEYTFPHLSLDQFC 591  
 500 AALYVVEGLEIEPLACPLVYEKTKRSMELKQAGHISHLMKRFPLRELVSEDRPRLPV 559  
 592 AALYVVEGLEIEPLACPLVYEKTKRSMELKQAGHISHLMKRFPLRELVSEDRPRLPV 651  
 560 LLAGCPVPLGVKQKLHWSLIGQOPNATTPGDTLDAFCLFETODKEVRLALNSFOGVW 619  
 652 LLAGCPVPLGVKQKLHWSLIGQOPNATTPGDTLDAFCLFETODKEVRLALNSFOGVW 711  
 620 LPINOMDLIASPFLQHCPLRKIVDVKGI FPRDESAAACPVPLMWRDKTLIEBQWE 679  
 712 LPINOMDLIASPFLQHCPLRKIVDVKGI FPRDESAAACPVPLMWRDKTLIEBQWE 771  
 680 DFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPCKIOTLMFRNAQTTPGVQHLMR 739

DB 772 DFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPCKIOTLMFRNAQTTPGVQHLMR 831  
 QY 740 IWMANRILRSINLGGTHLKEEDVMAACEALKHPKCLLESRLDCCGLTHACYLKISQILT 799  
 DB 832 IWMANRILRSINLGGTHLKEEDVMAACEALKHPKCLLESRLDCCGLTHACYLKISQILT 891  
 QY 800 TSPSLKSLSLAGNVTOGVWPLSDALRVSCALQKILPEDCGITTAGCCSLASALVSNR 859  
 DB 892 TSPSLKSLSLAGNVTOGVWPLSDALRVSCALQKILPEDCGITTAGCCSLASALVSNR 951  
 QY 860 SLTHCLSNNSLAGEVNLICRSWRLPHCSLQRLMLNQCCHDITAGCFALALAGNSWLT 919  
 DB 952 SLTHCLSNNSLAGEVNLICRSWRLPHCSLQRLMLNQCCHDITAGCFALALAGNSWLT 1011  
 QY 920 HSLSNMPVEDNGVYKLLCEVWREPSCHLODELVKCHLTAACCSLSCVISRSHLSLD 979  
 DB 1012 HSLSNMPVEDNGVYKLLCEVWREPSCHLODELVKCHLTAACCSLSCVISRSHLSLD 1071  
 QY 980 LTDNALDGGVAAALCEGLKQNSVLTGLKACGLTSDCCALSLASCRHLSLNVQ 1039  
 DB 1072 LTDNALDGGVAAALCEGLKQNSVLTGLKACGLTSDCCALSLASCRHLSLNVQ 1131  
 QY 1040 NNFSPKGMKLCSAFACPTSNLQIIGLMKQYPVQIRKLEEVOLKPRVVIDGSMHSD 1099  
 DB 1132 NNFSPKGMKLCSAFACPTSNLQIIGLMKQYPVQIRKLEEVOLKPRVVIDGSMHSD 1191  
 QY 1100 EDDRH 1104  
 DB 1192 EDDRY 1196

RESULT 6  
 US-10-677-943-24  
 ; Sequence 24, Application US/10677943  
 ; Publication No. US20040072297A1

; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United States of America as  
 ; APPLICANT: represented by the Secretary of the Department of Health and  
 ; APPLICANT: Human Services  
 ; APPLICANT: Nelson, Lawrence  
 ; APPLICANT: Tong, Zhi-Bin  
 ; TITLE OF INVENTION: Human Gene Critical to Fertility  
 ; FILE REFERENCE: 4239-64785  
 ; CURRENT APPLICATION NUMBER: US/10/677,943  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: 60/241,510  
 ; PRIOR FILING DATE: 2000-10-18  
 ; PRIOR APPLICATION NUMBER: PCT/US02/09776  
 ; PRIOR FILING DATE: 2002-03-29  
 ; PRIOR APPLICATION NUMBER: PCT/US01/10981  
 ; PRIOR FILING DATE: 2001-04-04  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 24  
 ; LENGTH: 1200  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-677-943-24

Query Match 81.2%; Score 5741.5; DB 15; Length 1200;  
 Best Local Similarity 95.9%; Pred. No. 0;  
 Matches 1098; Conservative 3; Mismatches 3; Indels 41; Gaps 1;  
 1 MEGKSLTFSSYGLQWCLYELDKKEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 60  
 52 MEGKSLTFSSYGLQWCLYELDKKEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 111  
 61 LHHEYGASLAWATSIISIFENMNLRTLSEKARDMKK----- 97  
 112 LHHEYGASLAWATSIISIFENMNLRTLSEKARDMKKSHSPDEPATWTDGPKSEKVPGI 171  
 98 -----ISQAMEBEGATAETBEQSIQAMEBEGATAETBEQSGCD 139

```
Db 172 SOAVQODSATAETKEOEISQAMEQEGATAETEOEISQAMEQEGATAETEOEGHGD 231
Qy 140 TWDKSHVMTKFAEBEEDVRSPENTADWPEMOTLAGAFSDRMGFRPRTVVLHGSGIG 199
Db 232 TWDKSHVMTKFAEBEEDVRSPENTADWPEMOTLAGAFSDRMGFRPRTVVLHGSGIG 291
Qy 200 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMOQKKESSVTEFISREMPDQAPTYEIM 259
Db 292 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMOQKKESSVTEFISREMPDQAPTYEIM 351
Qy 260 SRPERLLFIIDGPDLDGSLVNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 319
Db 352 SRPERLLFIIDGPDLDGSLVNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 411
Qy 320 VGEKLSSEVSPRYLLVRGISGEORJHLLEERGIGEHQKQGRALIMNNRELLDOCV 379
Db 412 VGEKLSSEVSPRYLLVRGISGEORJHLLEERGIGEHQKQGRALIMNNRELLDOCV 471
Qy 380 AVGSLICVALQLODVGESVAPFNQTLTGJHAAVFNHQLTPRGVRRCLNLEERVVKRF 439
Db 472 AVGSLICVALQLODVGESVAPFNQTLTGJHAAVFNHQLTPRGVRRCLNLEERVVKRF 531
Qy 440 CRMAVEGVNKRKSVFDGDDLMVOGLGSEELRALFHMNILLPDSHCEEYTFPHLSLODFC 499
Db 532 CRMAVEGVNKRKSVFDGDDLMVOGLGSEELRALFHMNILLPDSHCEEYTFPHLSLODFC 591
Qy 500 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSHLMKRFGLVSEEDVRPRLEV 559
Db 592 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSHLMKRFGLVSEEDVRPRLEV 651
Qy 560 LAGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCLFETODKEFVALANSFOEVM 619
Db 652 LAGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCLFETODKEFVALANSFOEVM 711
Qy 620 LPINQNDLIASSFCLOHCPYLKIRIVDVKGIFFPDESABACPVPLMMDKTLIEBQWE 679
Db 712 LPINQNDLIASSFCLOHCPYLKIRIVDVKGIFFPDESABACPVPLMMDKTLIEBQWE 771
Qy 680 DFCGMLGTHPHLRQLOLDGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLMR 739
Db 772 DFCGMLGTHPHLRQLOLDGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLMR 831
Qy 740 IVMANRNLRSNLGTHLKEEDVMAACALGHPKCLLESRLDCCGTLTHACYLKISQILT 799
Db 832 IVMANRNLRSNLGTHLKEEDVMAACALGHPKCLLESRLDCCGTLTHACYLKISQILT 891
Qy 800 TSPBLKSLISLAGNRYTDOGVMPISDALRVSOCAIQKLILEDGITTATGCGSLASALVSNR 859
Db 892 TSPBLKSLISLAGNRYTDOGVMPISDALRVSOCAIQKLILEDGITTATGCGSLASALVSNR 951
Qy 860 SLTTHLCJSSNLSGEBGVNLCRSMBLPHCSLQRLMLNCHDITAGCGTLALALMGNSMLT 919
Db 952 SLTTHLCJSSNLSGEBGVNLCRSMBLPHCSLQRLMLNCHDITAGCGTLALALMGNSMLT 1011
Qy 920 HLSTSNMPVEDNGVYKLCVWRPESCHLODELEVKCHLTAACCSLSVJSRSHLKSID 979
Db 1012 HLSTSNMPVEDNGVYKLCVWRPESCHLODELEVKCHLTAACCSLSVJSRSHLKSID 1071
Qy 980 LITDNLADGVAALCEGILKQNSVYTRIGLKAAGLTSDCCEALSALSCNHLTSLNLVQ 1039
Db 1072 LITDNLADGVAALCEGILKQNSVYTRIGLKAAGLTSDCCEALSALSCNHLTSLNLVQ 1131
Qy 1040 NNFPKGMKLCGSAFACPTSNLQITIGLMKQYPVQIRKLEBVOLKPRVVIDGSMHSFD 1099
Db 1132 NNFPKGMKLCGSAFACPTSNLQITIGLMKQYPVQIRKLEBVOLKPRVVIDGSMHSFD 1191
Qy 1100 EDDRH 1104
Db 1192 EDDRH 1196
```

```
us-10-860-761-4
; Sequence 4, Application US/10860761
; Publication No. US20040248775A1
; GENERAL INFORMATION:
; APPLICANT: WYETH
; TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS
; FILE REFERENCE: AM101318
; CURRENT APPLICATION NUMBER: US/10/860,761
; CURRENT FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1200
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-860-761-4
```

Query Match 81.2%; Score 5741.5; DB 16; Length 1200;

Best Local Similarity 95.9%; Pred. No. 0;

Matches 1098; Conservative 3; Mismatches 3; Indels 41; Gaps 1;

```
Qy 1 MEGDKSLTFSSYGLQWCLVELDKBEPTFKELKKSSSESTGSIPOFEIENANVECTAL 60
Db 52 MEGDKSLTFSSYGLQWCLVELDKBEPTFKELKKSSSESTGSIPOFEIENANVECTAL 111
Qy 61 LHEHYGASLAWATSSISIFENNMLRTSEKARDMMK----- 97
Db 112 LHEHYGASLAWATSSISIFENNMLRTSEKARDMMK----- 171
Qy 98 -----ISQAMEQEGATAETKEOEISQAMEQEGATAETEOEISQAMEQEGATAETEOEGHGD 139
Db 172 SOAVQODSATAETKEOEISQAMEQEGATAETEOEISQAMEQEGATAETEOEGHGD 231
Qy 140 TWDKSHVMTKFAEBEEDVRSPENTADWPEMOTLAGAFSDRMGFRPRTVVLHGSGIG 199
Db 232 TWDKSHVMTKFAEBEEDVRSPENTADWPEMOTLAGAFSDRMGFRPRTVVLHGSGIG 291
Qy 200 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMOQKKESSVTEFISREMPDQAPTYEIM 259
Db 292 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMOQKKESSVTEFISREMPDQAPTYEIM 351
Qy 260 SRPERLLFIIDGPDLDGSLVNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 319
Db 352 SRPERLLFIIDGPDLDGSLVNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 411
Qy 320 VGEKLSSEVSPRYLLVRGISGEORJHLLEERGIGEHQKQGRALIMNNRELLDOCV 379
Db 412 VGEKLSSEVSPRYLLVRGISGEORJHLLEERGIGEHQKQGRALIMNNRELLDOCV 471
Qy 380 AVGSLICVALQLODVGESVAPFNQTLTGJHAAVFNHQLTPRGVRRCLNLEERVVKRF 439
Db 472 AVGSLICVALQLODVGESVAPFNQTLTGJHAAVFNHQLTPRGVRRCLNLEERVVKRF 531
Qy 440 CRMAVEGVNKRKSVFDGDDLMVOGLGSEELRALFHMNILLPDSHCEEYTFPHLSLODFC 499
Db 532 CRMAVEGVNKRKSVFDGDDLMVOGLGSEELRALFHMNILLPDSHCEEYTFPHLSLODFC 591
Qy 500 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSHLMKRFGLVSEEDVRPRLEV 559
Db 592 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSHLMKRFGLVSEEDVRPRLEV 651
Qy 560 LAGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCLFETODKEFVALANSFOEVM 619
Db 652 LAGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCLFETODKEFVALANSFOEVM 711
Qy 620 LPINQNDLIASSFCLOHCPYLKIRIVDVKGIFFPDESABACPVPLMMDKTLIEBQWE 679
Db 712 LPINQNDLIASSFCLOHCPYLKIRIVDVKGIFFPDESABACPVPLMMDKTLIEBQWE 771
Qy 680 DFCGMLGTHPHLRQLOLDGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLMR 739
Db 772 DFCGMLGTHPHLRQLOLDGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLMR 831
```

QY 740 IVMANRNLRSNLGGTTHKEEDVRMACALKHPKCLLESRLDCCGLTHACTYKISQILT 799  
DB 832 IVMANRNLRSNLGGTTHKEEDVRMACALKHPKCLLESRLDCCGLTHACTYKISQILT 891  
QY 800 TSPSLKSLSLAGNKVTDOGVMPDSDALRVSOQALOKLILDECGITATGCSQSLASLVNRR 859  
DB 892 TSPSLKSLSLAGNKVTDOGVMPDSDALRVSOQALOKLILDECGITATGCSQSLASLVNRR 951  
QY 860 SLTHCLSNNSLGNNGVNLCSMRPLPHCSLORLMLNOCHLDTAGCGFLALAMNSWLT 919  
DB 952 SLTHCLSNNSLGNNGVNLCSMRPLPHCSLORLMLNOCHLDTAGCGFLALAMNSWLT 1011  
QY 920 HLSISMNVEVDNGVTLCEVWRBPSCHLODELVKCHLTAACCESTSCVTSRRLKSLD 979  
DB 1012 HLSISMNVEVDNGVTLCEVWRBPSCHLODELVKCHLTAACCESTSCVTSRRLKSLD 1071  
QY 980 LTDNALGDGVAALCEGLKOKNSVLTSLGLNACGLTSDCCALSLASCNRLTSLNLVQ 1039  
DB 1072 LTDNALGDGVAALCEGLKOKNSVLTSLGLNACGLTSDCCALSLASCNRLTSLNLVQ 1131  
QY 1040 NNFSPKGMKLCSPACPTSNLQITGLMKQYVQVIRKLEEVQLKRPVVIDGSMHSPD 1099  
DB 1132 NNFSPKGMKLCSPACPTSNLQITGLMKQYVQVIRKLEEVQLKRPVVIDGSMHSPD 1191  
QY 1100 EDDRH 1104  
DB 1192 EDDRH 1196

## RESULT 8

US-10-216-645-4  
; Sequence 4, Application US/10216645  
; Publication No. US20030125282A1  
; GENERAL INFORMATION:  
; APPLICANT: WEISS, BERTRAM  
; APPLICANT: LESSL, MONIKA  
; APPLICANT: PETERS-KOTTIG, MICHAELE  
; APPLICANT: BECKMANN, GEORG  
; TITLE OF INVENTION: HUMAN WATER PROTEINS  
; FILE REFERENCE: SCH-1910  
; CURRENT APPLICATION NUMBER: US/10216,645  
; CURRENT FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-216-645-4

Query Match 80.3%; Score 5683.5; DB 14; Length 1143;  
Best Local Similarity 95.4%; Pred. No. 0;  
Matches 1087; Conservative 5; Mismatches 12; Indels 35; Gaps 2;

QY 1 MEGDKSLTFSSYGLQWCLYEIDKEEFQTFKELLKKKSSSESTTCSIPQFIENAVETCAL 60  
DB 1 MEGDKSLTFSSYGLQWCLYEIDKEEFQTFKELLKKKSSSESTTCSIPQFIENAVETCAL 60  
QY 61 LLHEYYGASLMAWTSISIFENNMLTTLSEKARDMKKISQ-----AMEBEGATAATEEQ- 115  
DB 61 LLHEYYGASLMAWTSISIFENNMLTTLSEKARDMKKISQ-----AMEBEGATAATEEQ- 115  
QY 116 -----EISQAMEBEGATAATEEQAGGDTWMDYS 145  
DB 121 YGMTLLIGVSDISDSNNKHKTVGTHSSPABISQAMEBEGATAATEEQAGGDTWMDYS 180  
QY 146 HVMTFAAEEDVRASFENTADWPMQTLAAGFSDRWGFRPRVYLHGKSGIGKSALAR 205  
DB 181 HVMTFAAEEDVRASFENTADWPMQTLAAGFSDRWGFRPRVYLHGKSGIGKSALAR 240  
QY 206 RIVLCMAGGLYQGMFSTVFLPVREMRKSSSTETISREMPDSQAPVTEIMSRPRL 265

DB 241 RIVLCMAGGLYQGMFSTVFLPVREMRKSSSTETISREMPDSQAPVTEIMSRPRL 300  
QY 266 LFIIDGFDLGSVANNPTKLCCKNAEKOPETTLIRSLIRKTLPLPSFLITVRDVGTEKL 325  
DB 301 LFIIDGFDLGSVANNPTKLCCKNAEKOPETTLIRSLIRKTLPLPSFLITVRDVGTEKL 360  
QY 326 KSEVVSPPYLLVRISGQRTHILLERIGIHOHTQGLRAIMNRRELLDOCVAVGSLI 385  
DB 361 KSEVVSPPYLLVRISGQRTHILLERIGIHOHTQGLRAIMNRRELLDOCVAVGSLI 420  
QY 386 CVALQLODVESVAPFNQTLTGLHAAVFVHQTLPRGVRRCNLIEERVLAKRCMAVE 445  
DB 421 CVALQLODVESVAPFNQTLTGLHAAVFVHQTLPRGVRRCNLIEERVLAKRCMAVE 480  
QY 446 GVMNRKSVFDGDDIMVQGLSESELRALFHMNILLPSHCEERYTFPHLSLODFCALYYV 505  
DB 481 GVMNRKSVFDGDDIMVQGLSESELRALFHMNILLPSHCEERYTFPHLSLODFCALYYV 540  
QY 506 LEGEIEBPALCPLYVEKTSRMEKQAGFHHSLMMKRFPLGLVSEVRRPLFVTLGCPV 565  
DB 541 LEGEIEBPALCPLYVEKTSRMEKQAGFHHSLMMKRFPLGLVSEVRRPLFVTLGCPV 600  
QY 566 PLGVKQKLLHWSVSLGQOPNATTPGDTLDAFHCLEFETODKEFVRLALNSFOEVLPLINQ 625  
DB 601 PLGVKQKLLHWSVSLGQOPNATTPGDTLDAFHCLEFETODKEFVRLALNSFOEVLPLINQ 660  
QY 626 LDLIASSFCLOHCPYLRKIRVDVKGIFPRDSABACPVPLMMDKTLIEBQWDFCSML 685  
DB 661 LDLIASSFCLOHCPYLRKIRVDVKGIFPRDSABACPVPLMMDKTLIEBQWDFCSML 720  
QY 686 GTHPHLRQDLGSSILTERBANKTICAKLRHTCKIOTLMPNNAOITTEGVCHLWRIWVNR 745  
DB 721 GTHPHLRQDLGSSILTERBANKTICAKLRHTCKIOTLMPNNAOITTEGVCHLWRIWVNR 780  
QY 746 NLRSLNLGGTTHKEEDVRMACALKHPKCLLESRLDCCGLTHACTYKISQILTSPSLK 805  
DB 781 NLRSLNLGGTTHKEEDVRMACALKHPKCLLESRLDCCGLTHACTYKISQILTSPSLK 840  
QY 806 SLSLAGNKVTDOGVMPDSDALRVSOQALOKLILDECGITATGCSQSLASLVNRRSLTHLC 865  
DB 841 SLSLAGNKVTDOGVMPDSDALRVSOQALOKLILDECGITATGCSQSLASLVNRRSLTHLC 900  
QY 866 LSNNSLGNNGVNLCSMRPLPHCSLORLMLNOCHLDTAGCGFLALAMNSWLTSLISM 925  
DB 901 LSNNSLGNNGVNLCSMRPLPHCSLORLMLNOCHLDTAGCGFLALAMNSWLTSLISM 960  
QY 926 NPVEDNGVTLCEVWRBPSCHLODELVKCHLTAACCESTSCVTSRRLKSLDLDNAL 985  
DB 961 NPVEDNGVTLCEVWRBPSCHLODELVKCHLTAACCESTSCVTSRRLKSLDLDNAL 1020  
QY 986 GDGVAALCEGLKOKNSVLTSLGLNACGLTSDCCALSLASCNRLTSLNLVQNNFSPK 1045  
DB 1021 GDGVAALCEGLKOKNSVLTSLGLNACGLTSDCCALSLASCNRLTSLNLVQNNFSPK 1080  
QY 1046 GMMKLCSPACPTSNLQITGLMKQYVQVIRKLEEVQLKRPVVIDGSMHSPEDDRH 1104  
DB 1081 GMMKLCSPACPTSNLQITGLMKQYVQVIRKLEEVQLKRPVVIDGSMHSPEDDRH 1139

## RESULT 9

US-10-092-900A-348  
; Sequence 348, Application US/10092900A  
; Publication No. US20040043382A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Taudier Jr., Raymond J.  
; APPLICANT: Pena, Carol E.A.  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Gusev, Vladimir Y.

```
APPLICANT: Ji, Weizhen
APPLICANT: Gorman, Linda
APPLICANT: Miller, Charles E.
APPLICANT: Kekuda, Ramesh
APPLICANT: Patlurajan, Meera
APPLICANT: Gangoli, Esna A.
APPLICANT: Verneet, Corine A.M.
APPLICANT: Guo, Xiaojia Saaha
APPLICANT: Tchernev, Velizar T.
APPLICANT: Fernandes, Elma R.
APPLICANT: Caeman, Stacie J.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Gerlach, Valerie
APPLICANT: Liu, Yi
APPLICANT: Anderson, David W.
APPLICANT: Spaderma, Steven K.
APPLICANT: Catterton, Elina
APPLICANT: Leite, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alebrock, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Bugess, Catherine E.
FILE OF INVENTION: No. US20040043382a1e1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
PRIOR FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USSN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USSN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USSN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USSN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USSN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USSN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USSN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 348
LENGTH: 1033
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-900A-348

Query Match      72.3% Score 5113.5; DB 15; Length 1033;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 976; Conservative 7; Mismatches 18; Indels 21; Gaps 2;
```

```
Db 188 ERLFIIDGFDLGSVLANNDTKLCKDMAEKOPFTLIRSLRKVLPESEFLIVTRDVG 247
Qy 323 EKLSEVSPRYLLNGISGEORIHLLERGISEHOKTOGLRAIMNNRELLDOCVAVG 382
Db 248 EKSEVSPRYLLNGISGEORIHLLERGISEHOKTOGLRAIMNNRELLDOCVAVG 307
Qy 383 SLICVALQDQVGESEVAPNPQTLTGHAAFVHQTLPRGVARCLNEERVLKRCRM 442
Db 308 SLICVALQDQVGESEVAPNPQTLTGHAAFVHQTLPRGVARCLNEERVLKRCRM 367
Qy 443 AVEGWNKRSVFPDGDMLVQGLSEELRALFHNWILLPSHCERYTFPHSLDPCAL 502
Db 368 AVEGWNKRSVFPDGDMLVQGLSEELRALFHNWILLPSHCERYTFPHSLDPCAL 427
Qy 503 YVVLGELEIPALCPVLVEKTKRSMELKQGFPHISLMRPFGLVSEVRPPEVLG 562
Db 428 YVVLGELEIPALCPVLVEKTKRSMELKQGFPHISLMRPFGLVSEVRPPEVLG 487
Qy 563 CPVPLGVKQKLHWVSILGOOPNATTPGDTLDAFHCLFETQDKFVRLALNSFOEVLPI 622
Db 488 CPVPLGVKQKLHWVSILGOOPNATTPGDTLDAFHCLFETQDKFVRLALNSFOEVLPI 547
Qy 623 NONIDLIASSFCIQHCPYLKIRVDVKGIRPDEBSAECVPLWNRDKTLIEBQWDFC 682
Db 548 NONIDLIASSFCIQHCPYLKIRVDVKGIRPDEBSAECVPLWNRDKTLIEBQWDFC 607
Qy 683 SMGTHPHLQDLGSSILTERAMKTLCAKLRHPTCKIOTLMPRNAQITPGVOHLIRVM 742
Db 608 SMGTHPHLQDLGSSILTERAMKTLCAKLRHPTCKIOTLMPRNAQITPGVOHLIRVM 667
Qy 743 ANRLRSINLGTHLKEEDVRMACEALKHPKCLLESRLDCCGLYHACYLKISQILTSP 802
Db 668 ANRLRSINLGTHLKEEDVRMACEALKHPKCLLESRLDCCGLYHACYLKISQILTSP 727
Qy 803 SLKSLSLAGNKVTDQGMPLSDLRVSOCLQKYLIEDCGITATGCCSLASALVNSRLT 862
Db 728 SLKSLSLAGNKVTDQGMPLSDLRVSOCLQKYLIEDCGITATGCCSLASALVNSRLT 787
Qy 863 HLCISNNSLNEBGNILCRSMRLPHCSLORLMMNOCHLDPRAGGFLALAMGNSWLTHLS 922
Db 788 HLCISNNSLNEBGNILCRSMRLPHCSLORLMMNOCHLDPRAGGFLALAMGNSWLTHLS 847
Qy 923 LSNMPEVDNGVKLCEVWREPSCHLDLELVKCHLTAACCESSLCVLSRSHLKSJDLTD 982
Db 848 LSNMPEVDNGVKLCEVWREPSCHLDLELVKCHLTAACCESSLCVLSRSHLKSJDLTD 907
Qy 983 NALGDGVAAALCEGLKOKNSVTLRLGKACGLTSDCEALSLALSCNRHLTSLNLYONNF 1042
Db 908 NALGDGVAAALCEGLKOKNSVTLRLGKACGLTSDCEALSLALSCNRHLTSLNLYONNF 967
Qy 1043 SPKGMMLGSAFACPTSNLQIIGLMKQYVPQIRKLEEVQALKPRVVDGSMHSPDEDD 1102
Db 968 SPKGMMLGSAFACPTSNLQIIGLMKQYVPQIRKLEEVQALKPRVVDGSMHSPDEDD 1027
Qy 1103 RH 1104
Db 1028 RV 1029

RESULT 10
US-10-407-866-96
Sequence 96, Application US/10407866
Publication No. US20040002593A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
FILE REFERENCE: 66654-10(LJ 5755)
CURRENT FILING DATE: US/10/407,866
PRIOR APPLICATION NUMBER: US 60/370,538
PRIOR FILING DATE: 2002-04-04
```

NUMBER OF SEQ ID NOS: 129  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 96  
 LENGTH: 682  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-407-866-96

Query Match 46.9%; Score 3318.5; DB 15; Length 682;  
 Best Local Similarity 93.8%; Pred. No. 4.9e-274;  
 Matches 640; Conservative 1; Mismatches 0; Indels 41; Gaps 1;

QY 1 MEGDKSLTFSSYGLQWCLYEIDKEEFQTKELKKSSSESTCSIPQFENANVECLAL 60  
 DB 1 MEGDKSLTFSSYGLQWCLYEIDKEEFQTKELKKSSSESTCSIPQFENANVECLAL 60  
 QY 61 LHHEYGASLAWATSIIFENNNLRTLSKARDMDK----- 97  
 DB 61 LHHEYGASLAWATSIIFENNNLRTLSKARDMDKRSHPDPRATWTDQSPSKVPGI 120  
 QY 98 -----ISQAMEQEGATAETEOEISQAMEQEGATAETEOEISQAMEQEGATAETEOE 139  
 DB 121 SQAVQDSATPAETKEOEISQAMEQEGATAETEOEISQAMEQEGATAETEOE 180  
 QY 140 TWQYKSHVMTFAEBEDVRSFENTADMPMOTLAGAFSDRNGFRPTVVLHGKSGIG 199  
 DB 181 TWQYKSHVMTFAEBEDVRSFENTADMPMOTLAGAFSDRNGFRPTVVLHGKSGIG 240  
 QY 200 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMRKSSVTEFISREMPDSPA 259  
 DB 241 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMRKSSVTEFISREMPDSPA 300  
 QY 260 SRPRLFLITDGFDDGSLVNNDTCLCDMAEKOPPTLIRSLRKVLPPESFLIVYRD 319  
 DB 301 SRPRLFLITDGFDDGSLVNNDTCLCDMAEKOPPTLIRSLRKVLPPESFLIVYRD 360  
 QY 320 VTEKLSKSEVSPRYLVRSIGSORIHLLEERGIGEOQTKGLRAIMNNELDDCCVP 379  
 DB 361 VTEKLSKSEVSPRYLVRSIGSORIHLLEERGIGEOQTKGLRAIMNNELDDCCVP 420  
 QY 380 AVGLICVALQLODVGSVAPFNQTLTGHAAPFVHQLTPRGVRRCLNLEERVYLRK 439  
 DB 421 AVGLICVALQLODVGSVAPFNQTLTGHAAPFVHQLTPRGVRRCLNLEERVYLRK 480  
 QY 440 CRMAVEGVNRKSPFDGDLAVQGISSEELALTHANILLPDSCEERYTFPHSLDQFC 499  
 DB 481 CRMAVEGVNRKSPFDGDLAVQGISSEELALTHANILLPDSCEERYTFPHSLDQFC 540  
 QY 500 AALYYVLEGLIEPAPLGLYVEKTRSMELKQAGFHISLMKRFLRGLVEDVRRPLEV 559  
 DB 541 AALYYVLEGLIEPAPLGLYVEKTRSMELKQAGFHISLMKRFLRGLVEDVRRPLEV 600  
 QY 560 LLAGCPVPLGVKQKLHWSLLGQOPNATTPGDTLDAFHCLFETODKEFVALINSPOEW 619  
 DB 601 LLAGCPVPLGVKQKLHWSLLGQOPNATTPGDTLDAFHCLFETODKEFVALINSPOEW 660  
 QY 620 LPIQNIDLIIASFCLQHCPLYL 641  
 DB 661 LPIQNIDLIIASFCLQHCPLYL 682

RESULT 11  
 US-10-216-645-5  
 Sequence 5, Application US/10216645  
 Publication No. US20030125282A1  
 GENERAL INFORMATION:  
 APPLICANT: WEISS, BERTRAM  
 APPLICANT: LESSL, MONIKA  
 APPLICANT: PETERS-KOTTIG, MICHAEL  
 APPLICANT: BECKMANN, GEORG  
 TITLE OF INVENTION: HUMAN MATER PROTEINS  
 FILE REFERENCE: SCH-1910  
 CURRENT APPLICATION NUMBER: US/10/216.645

CURRENT FILING DATE: 2003-01-21  
 PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
 PRIOR FILING DATE: 2001-08-10  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 1111  
 TYPE: PRT  
 ORGANISM: Mus sp.  
 US-10-216-645-5

Query Match 38.0%; Score 2688; DB 14; Length 1111;  
 Best Local Similarity 51.6%; Pred. No. 8.8e-220;  
 Matches 544; Conservative 183; Mismatches 281; Indels 46; Gaps 10;

QY 74 TSISIFENNNLRTLSKARDMDKSIQ-----AMEQEGATAETEE 114  
 DB 77 STWSPSENVSRAIIKDSGEVEQASERKMTSPENDSKSIQKQGPPOEOTSEYLSKEE 136  
 QY 115 QEISQAMEQEGATAETEEQHGDTWDYKSHVMTFAEBEDVRSFENTADMPMOTL 174  
 DB 137 DEVEA-----DKDNGGDLQDYRAHVIATKFTSDVDLH-----YDSPEKLL 177  
 QY 175 AGAFSDRMGFRPTVVLHGKSGIGKALARRIVLCMAOGGLYQGMFSYVFLPVREMR 234  
 DB 178 SDAPKPYQKTPQPTIILHGRPGVKSALARSIVLCMAOGGLYQGMFSYVFLPVREMR 236  
 QY 235 KKESSVTEFISREMPDSPA 293  
 DB 237 TEKSLAQLLAKECPDSDMDLVTKIMSPERLLFVIDGLDMDSDVLAQDDMTLSRMDQEQ 296  
 QY 294 PPTLIRSLRKVLPPESFLIVYRDVTEKLSKSEVSPRYLVRSIGSORIHLLEERG 353  
 DB 297 PPTLIRSLRKVLPPESFLIVYRDVTEKLSKSEVSPRYLVRSIGSORIHLLEERG 356  
 QY 354 IGEHOKTQGLRAIMNNELDDCCVP 413  
 DB 357 SNBSDRLOVHSHSLIENQGLDQCAQPSVCSLVCEALDQKLGRCCTLPQTLTGATYL 416  
 QY 414 VFHQLTPRGVRRCLNLEERVYLRKFCMAVEGVNRKSPFDGDLAVQGISSEELALF 473  
 DB 417 VFHQLTPRGVRRCLNLEERVYLRKFCMAVEGVNRKSPFDGDLAVQGISSEELALF 476  
 QY 474 HNNILLPDSH-CEERYTFPHSLDQFCALYYVLEGL-ELPAPLGLYVEKTRSMELKQ 531  
 DB 477 HNNILLPDSH-CEERYTFPHSLDQFCALYYVLEGL-ELPAPLGLYVEKTRSMELKQ 534  
 QY 532 AGFHISLMKRFLRGLVEDVRRPLEVLCGPVPLGVKQKLHWSLLGQOPNATTPG 591  
 DB 535 TD-DTRLGMRFLPGLMNDILKTLVLEPYVLPVVEQDLQHWVSLAQOVNGTSPMD 593  
 QY 592 TLDAFHCLFETODKEFVALINSPOEWLPIQNIDLIIASFCLQHCPLYLRKIRVDYKGI 651  
 DB 594 TLDAFHCLFETODKEFVALINSPOEWLPIQNIDLIIASFCLQHCPLYLRKIRVDYKGI 653  
 QY 652 PPRDESADACPVVPLM-MRDKTLIEQWEDPCSMGLGTHPHRLDLCSSITLERAKTLC 710  
 DB 654 LSVYNTLELCPVVVYQETQCKPLIMWNGFCSVLGSLRLKEDLDLDSILISQAPMILC 713  
 QY 711 AKLHPQCKIOTMFRNAQITPGVHLMRIVMANRNLSNLGTHLKEEDVRRACALK 770  
 DB 714 LELNOSCRLOKLPFSAEVAVSGKHLMKLFSYNNKLYNLKTLGTPKDDMDKLCALCALK 773  
 QY 771 HPKCLLESRLDDCGSLTHACYLKSQILITPSPSIKSLSLAGNKVTDGVMPLPALRVSQ 830  
 DB 774 HKSCSVETLRDSEELIITIGEMISTILLITTHKCLSLAKNRYGVGWSMISLGNALSSM 833  
 QY 831 CALQKTLLEDCGTATGCGSLASALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHCSL 890  
 DB 834 CLQKTLLEDCGTATGCGSLASALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHCSL 893  
 QY 891 ORLMLNCHDITAGCGGLALALMGNSWLTHLSLSPNVEDNGVYLCEVWRBPSCHDIL 950

Db 894 QRLILNHCNIVDYAYGFLMARLANNTKLTLSLTNNPVGDMKLLCEALKEPTCYLOEL 953  
Qy 951 ELYKCHITTAACCSLSCVISRRLKSLDLTDNALGOGVNALEGLKOKNSVLTRLGLK 1010  
Db 954 ELVVCOLTONCEDLACMITTTKHLKSLDLGNNAAGDGVTLTLEGLKSSSRRRGLG 1013  
Qy 1011 ACGLTSDCCALSLATSCNRHLTSLNTVONNFSPKMKLCSAPACPTSNLOIIGLWKQ 1070  
Db 1014 ACTLNSCCALSLATSCNPHLNSLNTLVKNDFTSGMLKLCASAPCPVSNIGIIGLWKQ 1073  
Qy 1071 YPVOIRKLEEVOLKPRVVIDSGMHSFDEDDRH 1104  
Db 1074 YVARVRQLEBEVEFKPHVVIDGDWYASDEDDRN 1107  
RESULT 12  
US-10-399-443-6  
Sequence 6, Application US/10399443  
Publication No. US20040028669A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as Represented by the  
APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
APPLICANT: Health  
APPLICANT: Nelson, Lawrence M.  
APPLICANT: Tong, Zhi-Bin  
APPLICANT: Nelson, Lawrence  
APPLICANT: Zhi-Bin, Tong  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64785  
CURRENT APPLICATION NUMBER: US/10/399,443  
CURRENT FILING DATE: 2003-04-16  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6  
LENGTH: 1111  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-399-443-6  
Query Match 36.4%; Score 2577; DB 15; Length 1111;  
Best Local Similarity 50.8%; Pred. No. 2.7e-210;  
Matches 535; Conservative 183; Mismatches 290; Indels 46; Gaps 10;  
Qy 74 TSSIFENMMLRTLSEKARDMKKISQ-----AMEGGATPAETEE 114  
Db 77 STMSPESEVSRALIKDSGEVEEQASERKMTSPENDSKSIQKQGPQEQTSFTLQSKEB 136  
Qy 115 QEISQAMEGEGATAEFEQGHGSDTWDYKSHWTKRAEEDVRSFENTAAADPEKQTL 174  
Db 137 DEVEA-----DKONGGDLDQYKAHVIAKTDVSDLA-----YDSEPKKL 177  
Qy 175 AGAFSDRWGFRPRTVVLHGSGIGKALARIYLCAAGGLYOGMFYVFLPVRMOR 234  
Db 178 SDAKPYQKTPQPHITIIHGRPGVKSALASIVLGAAGLFGKM-SFVIFSVREIKT 236  
Qy 235 KKESSVTEFSREMPDSQAPVTEIMSRPERILFIIDGFDIGSVL-NNDTYKCDMAEK 293  
Db 237 TEKSSLAQLTAKECPDSTDVLTAKIMSQPERILFIIDGDMDSVLQHDMDTLSDTQDEQ 296  
Qy 294 PPFLLISLARKVLPESFLIVRVDTETLKSEVSPRLVLRGISGEGRIHLLERG 353  
Db 297 PYYLMLSLARKALLPQSFLLITTRNGLELTKSMVVSPLTIVLVEGSSASRSQVLENI 356  
Qy 354 IGEHQKQGLAIANNBEILDQCOVPAVSLICVALQLODVAGSVAPFNQTLTGLHAAAF 413  
Db 357 SNESDRIQVHSLLENHQLPQCCAPSVCSLVCAALQLOKLGKRCULPQCTTLGLYATL 416  
Qy 414 VFHQLTGRGVVRCINTLEERVVLKRCMAVEGVNKRKSVFDDGLMVGIGSESELRALP 473

Db 417 VFHQLTGRBPSOASLQEOBITLVGLCMMAEGVTNRSVFYDDDLKNVSLKSEILLALP 476  
Qy 474 HNNILPDSH-CERYYPFPLSLQDFPCALYYVLEGL-EYEPALCPLYVEKTRSMELKQ 531  
Db 477 HNNILLVGNHSECCVFSHLSLQDFPALYYVLEGLLEFNQHC-FIENQNSIMVYK 534  
Qy 532 AGPHISLMMKRPFLGLVSDVRRPLEVLLGCPVPLGKOKLHMVSLGQOPNATTPGD 591  
Db 535 TD-DTRLGKMRFLFGLMNNDILKTELEVLFEYPIVPEVQKQHTVSLAQOVGTSRMD 593  
Qy 592 TLDAFGLFETODEFVRALANSEFOEVLPIQNLDLIASSFCLOHCPYLRKTRVDYKGI 651  
Db 594 TLDAFYCLFESODEEFVGALEKFOEVTLLINQMDKLVSSYCKHQONKALRVDIRDL 653  
Qy 652 FPRDESAECPPVPLM-NRDKTLIEBQWEDFCSMLGTHPHLRQDLGSSITLERAMTLC 710  
Db 654 LSVNLTLECPVTVQETQCKPLMETTGNFCVSLSRNLKEILDGDSILSQRAMKILC 713  
Qy 711 AKLRHPYCKIQTLMFNAQITPVOHLMRTVMANRNLRSINTLGTHLKEEDVMACEALK 770  
Db 714 LEARNQSCRQKLTFSKAEVVSGLKHLTKLFSNQNLKYINTLGTPMKDDMKLACGALK 773  
Qy 771 HPKCLLESLRDOCCGLTHACYLKSQILTTSPSLKSLSLAGNKYTDGCVNPLDALRVQ 830  
Db 774 HPKCSVETRLDSCELTIIGYEMISTLLISTTRKCLSLAKNRGVKSMISLGNALSSM 833  
Qy 831 CALQKLEBQCGITATGCSLASALVNSRLTTHCLSNNSIGBGNVLLCRSMRLPHCSL 890  
Db 834 CLQKQLTLDNCGLTLPASCHLVSLFSNQNLTHCLSNNSLGTBGOVQLOQPLNBPCLAL 893  
Qy 891 QRLILNHCNIVDYAGCGFLATALMGNSVLTSLSMNPEVDNGVYLLCEVAREPSCHLODL 950  
Db 894 QRLILNHCNIVDYAYGFLMARLANNTKLTLSLTNNPVGDMKLLCEALKEPTCYLOEL 953  
Qy 951 ELYKCHITTAACCSLSCVISRRLKSLDLTDNALGOGVNALEGLKOKNSVLTRLGLK 1010  
Db 954 ELVVCOLTONCEDLACMITTTKHLKSLDLGNNAAGDGVTLTLEGLKSSSRRRGLG 1013  
Qy 1011 ACGLTSDCCALSLATSCNRHLTSLNTVONNFSPKMKLCSAPACPTSNLOIIGLWKQ 1070  
Db 1014 ACTLNSCCALSLATSCNPHLNSLNTLVKNDFTSGMLKLCASAPCPVSNIGIIGLWKQ 1073  
Qy 1071 YPVOIRKLEEVOLKPRVVIDSGMHSFDEDDRH 1104  
Db 1074 YVARVRQLEBEVEFKPHVVIDGDWYASDEDDRN 1107  
RESULT 13  
US-10-677-943-6  
Sequence 6, Application US/10677943  
Publication No. US20040072297A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as  
APPLICANT: represented by the Secretary of the Department of Health and  
APPLICANT: Human Services  
APPLICANT: Nelson, Lawrence  
APPLICANT: Tong, Zhi-Bin  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64790  
CURRENT APPLICATION NUMBER: US/10/677,943  
CURRENT FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US02/09776  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 6  
LENGTH: 1111  
TYPE: PRT  
ORGANISM: Mus musculus



US-10-677-943-6

Query Match 36.4%; Score 2577; DB 15; Length 1111;  
Best Local Similarity 50.8%; Pred. No. 2.7e-210;  
Matches 535; Conservative 183; Mismatches 290; Indels 46; Gaps 10;

QY 74 TSISIFENMNLRTLSKARDMMKISQ-----AMEOGATTAETEE 114  
DB 77 STMPSSEVSRALIKDSESEVEQASERKMTSPENDSKIQKQDQPEBOJSETIQSEE 136  
QY 115 QEISQAMEQEGATAETEEQGHGDDTWDYKSHVMTKFAEEDVRRSFENTADWPEWOTL 174  
DB 137 DEVEA-----DKONGGDLQDYKAHVIKFPDVSIDLH-----YDSEPMKLL 177  
QY 175 AGARDSRWKGRPRVTYHSGSGICKSALARRIVLCMAQSGLYQGMESYVFPFLPREKOR 234  
DB 178 SDAKRPYOKTQOPHTIILHGRPGVKSALARSIVLGTQGXLPQKM-SFVIFFSVREKXT 236  
QY 235 KKESSVTEFISREWPDSQAPYTEIMSRPERLFIIDGFDLGSYV-LNNDTKLCKDMAEKO 293  
DB 237 TEKSLAQILAKECPDSDTDLTKIMSOPERLIFVIGDDMDSDYLQHDMDTLSDPTQEQ 296  
QY 294 PPTLIRSLRKVLIPESFLIVYADVGTETKSESVSPRYLLVNGISGEORIHLLERG 353  
DB 297 PITYMYSLLARKALLPQSFLLITTRNTGLEKKSMMVSPVYLIVGSLASRSQVLVEMI 356  
QY 354 IGEHOKTQGLAIMNRRRLDQCOVPAGSLICVALQIQDVVGSVAFPNQTLTGHAAF 413  
DB 357 SNESDRIQVHSLIENHQLPQCOAPSVCSLVCEALQIQKGLKRCCTLPCQTLTGLVATL 416  
QY 414 VFHOLTFRGVVRCINLEERVVLKRCFMAVEGVNRRKSPDGDLLMWQGLGESELRALF 473  
DB 417 VFHOLTFRGPQSLAQSEBOITLVGLCMAABEGVTMRSVYDDLLKYSLSKESEIILAF 476  
QY 474 HNNILPDSH-CEBYTTFPHSLDQFCALYYVLEGL-EIEPALCPLYVEKTKSMELKO 531  
DB 477 HNNILQVGHNSQECYVFSHSLDQFALYYVLEGEETNQHC-FIENQRSIMEYKR 534  
QY 532 AGFIHSHMMKRFPLGVSEVVRPRLVLLCGPVLGVKQKLLMWSLGGQPNATTPGD 591  
DB 535 TD-DTRLGMRKPLFGMLNKDILKLTLELPEYVLPYTEQKLOHTVSLIAQOVNGTSPMD 593  
QY 592 TLDAFHCLETQDDEFPVRLAANSFOEVLPINONDLIASSFCLOHCHYLKIRIVYGI 651  
DB 594 TLDAFHCLETQDDEFPVRLAANSFOEVLPINONDLIASSFCLOHCHYLKIRIVYGI 653  
QY 652 FPRDESACAPVPLM-WRDKTLIEQWEDFCMSLGTPHLRDLGSSITTEBAMKTL 710  
DB 654 LSVNTELECPVTVVQETQCKPLMETTGNFCVLSIRNKEJDLGDSISQRAKILC 713  
QY 711 AKLRHPCKICTLMEFNAQITPGVOHLMRIYMANRNLISNLGTHLKEEDVRMACEALK 770  
DB 714 LELNOSORLOKLTFKSAEVVSGKLHLTKLLFSNONLKYLMGTPMKDMDKLAACEALK 773  
QY 771 HPKCLSLRSLCCGLTACVLYLSQILTSPLSKSLSLAGNXYTDCQVMPLSALARRSQ 830  
DB 774 HPKCSVETLRDSCBELTIGYEMISTLLSTRICKSLANRRGVMSMISLGNALSSSM 833  
QY 831 CALOKLILDECGITATGQCSLASALVNSRSLTHLCLSNNSLGNQVNLICRSMELPHCSL 890  
DB 834 CLQKLLINDNGLIPASCHLLVSALLFSNONLTHLCLSNNSLGTSGVOQLCOFLANPEAL 893  
QY 891 QRLMLNCHLDTAGCGFLALALMGNISWTLHSLSNPNVEDNGVLYLCEVMREPSCHLOL 950  
DB 894 QRLMLNCHLDTAGCGFLALALMGNISWTLHSLSNPNVEDNGVLYLCEVMREPSCHLOL 953  
QY 951 ELVYCHLTAAACCSLSCVSRHLKSLDLTDNALGQGVNALLCEGLKOKXSVLTRGLK 1010  
DB 954 ELVYCHLTAAACCSLSCVSRHLKSLDLTDNALGQGVNALLCEGLKOKXSVLTRGLK 1013  
QY 1011 ACGLTSDCCALSLALSNCNRHLSLNLVONNFSFKGMKLCSAFACPTSNQIIGLWMQ 1070  
DB 1014 ACGLTSDCCALSLALSNCNRHLSLNLVONNFSFKGMKLCSAFACPTSNQIIGLWMQ 1073

QY 1071 YPVOIRKLEEVOLLKRPVYIDSGWSHSPDEDRH 1104  
DB 1074 YPARVROLESEVFEKPHVYIDGTASDEDRN 1107

RESULT 14  
US-10-399-443-2  
; Sequence 2, Application US/10399443  
; Publication No. US20040028659A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: US/10/399,443  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 2  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-399-443-2

Query Match 27.4%; Score 1940; DB 15; Length 385;  
Best Local Similarity 98.7%; Pred. No. 1.4e-156;  
Matches 380; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 252 QAPYTEIMSRPERLFIIDGFDLGSYVLANNDTKLCKDMAEKOPFTLIRSLIRVLLPES 311  
DB 1 QAPYTEIMSRPERLFIIDGFDLGSYVLANNDTKLCKDMAEKOPFTLIRSLIRVLLPES 311  
QY 312 FLIYVADVGTETKSESVSPRYLLVNGISGEORIHLLERGIGEHOKTQGLRAIMNRE 371  
DB 61 FLIYVADVGTETKSESVSPRYLLVNGISGEORIHLLERGIGEHOKTQGLRAIMNRE 120  
QY 372 LLDQCOVPAGSLICVALQIQDVVGSVAFPNQTLTGHAAFVHOLTFRGVVRCINLE 431  
DB 121 LLDQCOVPAGSLICVALQIQDVVGSVAFPNQTLTGHAAFVHOLTFRGVVRCINLE 180  
QY 432 ERVVLKRCFMAVGVNRRKSPDGDLLMWQGLGESELRALFHHNILLPDSHCEYYTFF 491  
DB 181 ERVVLKRCFMAVGVNRRKSPDGDLLMWQGLGESELRALFHHNILLPDSHCEYYTFF 240  
QY 492 HLSIQDCCALYYVLEGELEIPALCPLYVEKTKRSMELKQAGFIHSHLMKRFGLVSE 551  
DB 241 HLSIQDCCALYYVLEGELEIPALCPLYVEKTKRSMELKQAGFIHSHLMKRFGLVSE 300  
QY 552 DVRRPFLVLLCGPVLGVKQKLLHVSLLGQPNATTPGDTLDAFHCLFETQDEFPVRLA 611  
DB 301 DVRRPFLVLLCGPVLGVKQKLLHVSLLGQPNATTPGDTLDAFHCLFETQDEFPVRLA 360  
QY 612 LNSFOEVLPINONDLIASSFCILQ 636  
DB 361 LNSFOEVLPINONDLIASSFCILQ 385

RESULT 15  
US-10-677-943-2  
; Sequence 2, Application US/10677943  
; Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as



; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zh1-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT APPLICATION NUMBER: US/10/677,943  
; CURRENT FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US02/09776  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 2  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-677-943-2

Query Match 27.4%; Score 1940; DB 15; Length 385;  
Best Local Similarity 98.7%; Pred. No. 1.4e-156;  
Matches 380; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 252 QAPTEIMSRPERLLFIIDGDDIGSVLNDTKLCKDMAEKQPPPTLIRSLIRKVLLES 311  
DB 1 QAPTEIMSRPERLLFIIDGDDIGSVLNDTKLCKDMAEKQPPPTLIRSLIRKVLLES 60  
QY 312 FLIYTVRDVGTETKSEVSPRYLLVRGISGEORIHLLERIGEGHOKTQGLRAIMNRE 371  
DB 61 FLIYTVRDVGTETKSEVSPRYLLVRGISGEORIHLLERIGEGHOKTQGLRAIMNRE 120  
QY 372 LLDCCQVPVAGSLICVALQLQDVVGEVAPFNQTLTGILHAAVFPHQLTPRGVARRCLNLE 431  
DB 121 LLDCCQVPVAGSLICVALQLQDVVGEVAPFNQTLTGILHAAVFPHQLTPRGVARRCLNLE 180  
QY 432 ERVVLKRFPCMAVGEVGNRRKSVFPGDDLMVQGLGESELRALFHMNILLPSHCEEYTF 491  
DB 181 ERVVLKRFPCMAVGEVGNRRKSVFPGDDLMVQGLGESELRALFHMNILLPSHCEEYTF 240  
QY 492 HLSIQDFCAALYYVLEGLIEIPALCPLYVEKTKRSMELKQAFHHSIMMKRFLFGVSE 551  
DB 241 HLSIQDFCAALYYVLEGLIEIPALCPLYVEKTKRSMELKQAFHHSIMMKRFLFGVSE 300  
QY 552 DVRRPLEVLLGCPVPLGVKQKLHMSVSLGQQPNATTGGDTLDAFHCLFETQDKEFVRLA 611  
DB 301 DVRRPLEVLLGCPVPLGVKQKLHMSVSLGQQPNATTGGDTLDAFHCLFETQDKEFVRLA 360  
QY 612 LNSFOEYWLPIINQNLDIASSFCIQ 636  
DB 361 LNSFOEYWLPIINQNLDIASSFCIQ 385

Search completed: July 19, 2005, 12:19:11  
Job time : 479 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 22:13:46 ; Search time 17004 Seconds  
(without alignments)  
11498.283 Million cell updates/sec

Title: US-10-066-521-5  
Perfect score: 4035  
Sequence: 1 atggaaggagacaatcgct.....tcgaggggcgtgtctttaa 4035

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hhg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_rc:\*  
11: gb\_rgs:\*  
12: gb\_ry:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2917	72.3	3489	6	AX478549 Sequence
2	2917	72.3	3485	6	AY154460 Homo sapi
3	2917	72.3	3926	6	AX704821 Sequence
4	2862	70.9	3830	6	AX704823 Sequence
5	2803	69.5	5859	6	AX459881 Sequence
6	2803	69.5	6939	6	AX459873 Sequence
7	2803	69.5	6939	6	AX459891 Sequence
8	2713	67.2	3885	6	AY054986 Homo sapi
9	2713	67.2	3900	6	AX427610 Sequence
10	1936	48.0	2753	6	CQ731113 Sequence
11	1549	38.4	157141	9	AC011470 Homo sapi
12	1549	38.4	193609	9	AC024580 Homo sapi
13	1102	27.3	1157	6	AX427588 Sequence
14	900	22.3	167509	2	AC012107 Homo sapi
15	657	16.3	1075	6	AX427590 Sequence
16	509	12.6	167509	2	AC012107 Homo sapi
17	219	5.4	164824	2	AC023887 Homo sapi
18	174	4.3	164824	2	AC023887 Homo sapi
19	165	4.1	1705	9	BC040925 Homo sapi

c	20	24	0.6	24	6	AX427600 Sequence
c	21	24	0.6	24	6	AX427601 Sequence
c	22	23	0.6	439	4	BR00C0RF Z66039 B. taurus MR
c	23	23	0.6	3545	4	AY721594 Bos tauru
c	24	23	0.6	159506	9	HS3418 AL021918 Human DNA
c	25	23	0.6	247877	2	AC096265 Rattus no
c	26	23	0.6	307855	2	AC134117 Rattus no
c	27	22	0.5	22	6	AX427594 Sequence
c	28	22	0.5	26	6	AX427596 Sequence
c	29	22	0.5	27	6	AX427604 Sequence
c	30	22	0.5	589	11	BV161832 RPAMSE00
c	31	22	0.5	1576	10	FI43559507 AF143565 Mus muscu
c	32	22	0.5	3303	10	AY329488 Mus muscu
c	33	22	0.5	3303	10	AY329491 Mus muscu
c	34	22	0.5	3372	10	AY329485 Mus muscu
c	35	22	0.5	3375	10	AY329489 Mus muscu
c	36	22	0.5	3405	10	AY329487 Mus muscu
c	37	22	0.5	3432	10	AY329484 Mus muscu
c	38	22	0.5	3447	6	AX427592 Sequence
c	39	22	0.5	3447	10	AF074018 Mus muscu
c	40	22	0.5	3453	10	AY329486 Mus muscu
c	41	22	0.5	3454	10	AY329490 Mus muscu
c	42	22	0.5	3470	10	AY196362 Mus muscu
c	43	22	0.5	3480	10	AY196361 Mus muscu
c	44	22	0.5	3534	10	BC053384 Mus muscu
c	45	22	0.5	55310	2	AC102436 Mus muscu

#### ALIGNMENTS

RESULT 1  
AX478549  
LOCUS AX478549 3489 bp DNA linear PAT 12-AUG-2002  
DEFINITION Sequence 3 from Patent WO0248362.  
ACCESSION AX478549  
VERSION AX478549.1 GI:22217318  
KEYWORDS

SOURCE  
ORGANISM Homo sapiens (human)

REFERENCE  
1 Rankumar, J. and Arvizu, C.  
Embryogenesis associated proteins  
Mammalia; Eutheria; Primates; Cetartihini; Homnidae; Homo.

AUTHORS  
TITLE  
JOURNAL  
Patent: WO 0248362-A 3 20-JUN-2002;  
Incyte Genomics, Inc. (US)

FEATURES  
source  
1..3489  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7474830CB1"

#### ORIGIN

Query Match 72.3%; Score 2917; DB 6; Length 3489;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 290 AAATTTCAAGCTATGGAACAAGAGTGCACAGCAGCAGACAGACAGACAGACAGAA 349  
452 AAATTTCAAGCTATGGAACAAGAGTGCACAGCAGCAGACAGACAGACAGACAGAA 511  
QY 350 TTTTCAAGCTATGGAACAAGAGTGCACAGCAGCAGACAGACAGACAGACAGACATG 409  
512 TTTTCAAGCTATGGAACAAGAGTGCACAGCAGCAGACAGACAGACAGACAGACATG 571  
DB 410 GAGGTGACACATGGACATCAAGAGTCACTGATGATCACTAAATTGCTGAGAGAGATG 469  
572 GAGGTGACACATGGACATCAAGAGTCACTGATGATCACTAAATTGCTGAGAGAGATG 631  
QY 470 TACGCTAGATTGTAANAACCTGCTGATGCTGACCGGGAATGCAAACTGGCTGCTG 529  
|||||

Db 632 TACGCTGATGTTTAAAAACACTGCTGCTGACTGGCCGGAATGCAAACTGTGGCTG 691  
Qy 530 CTTTGTATTCAGACCGGTGGGGCTTCGGCCCTCGACGGGTGTCTGACGGAAAGTGA 589  
Db 682 CTTTGTATTCAGACCGGTGGGGCTTCGGCCCTCGACGGGTGTCTGACGGAAAGTGA 751  
Qy 590 GAATTTGGGAAATTCGGCTTGAAGCAGAGAGATCGTCTGTGCTGGCGCAGAGTGA 649  
Db 752 GAATTTGGGAAATTCGGCTTGAAGCAGAGAGATCGTCTGTGCTGGCGCAGAGTGA 811  
Qy 650 ACCAGGAAATGTTCTCTACGCTCTTCTCTCCCGCTTGAAGAGATGACGGAAAG 709  
Db 812 ACCAGGAAATGTTCTCTACGCTCTTCTCTCCCGCTTGAAGAGATGACGGAAAG 871  
Qy 710 AGAGAGGTGTCAGAGTTCATCTCAGAGAGTGGCCGACCTCCAGGCTCCGGTGA 769  
Db 872 AGAGAGGTGTCAGAGTTCATCTCAGAGAGTGGCCGACCTCCAGGCTCCGGTGA 931  
Qy 770 AGATCATGTCCCGACCAAGAAAGCTGTGTTTCATGATGACGGTTCGATGACCTGG 829  
Db 932 AGATCATGTCCCGACCAAGAAAGCTGTGTTTCATGATGACGGTTCGATGACCTGG 991  
Qy 830 CTGTCTCTCAATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAAAGCCTCCGTTCA 889  
Db 992 CTGTCTCTCAATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAAAGCCTCCGTTCA 1051  
Qy 890 CCTCTATGCGAGTGTGCTGAGAGAGGTCTGCTCCGATGCTTCCGATCTGACCG 949  
Db 1052 CCTCTATGCGAGTGTGCTGAGAGAGGTCTGCTCCGATGCTTCCGATCTGACCG 1111  
Qy 950 TCAGAGAGTGGGACAGAGAAAGCTCAAGTCAGAGGTGTGTCTCCGTTACCTGTTAG 1009  
Db 1112 TCAGAGAGTGGGACAGAGAAAGCTCAAGTCAGAGGTGTGTCTCCGTTACCTGTTAG 1171  
Qy 1010 TTAGAGAAATCTCCGGGGAAACAAAGATTCATCTCTCTTGAAGCGCGGATTTGTAG 1069  
Db 1172 TTAGAGAAATCTCCGGGGAAACAAAGATTCATCTCTCTTGAAGCGCGGATTTGTAG 1231  
Qy 1070 ATCAAGAGACAAAGGTGTGCTGATCAATGAACCGTGAAGTGTGCGACAGTGC 1129  
Db 1232 ATCAAGAGACAAAGGTGTGCTGATCAATGAACCGTGAAGTGTGCGACAGTGC 1291  
Qy 1130 AGGTCCCGCGGTGGGCTCTCTCATCTGCTGCTGCTGAGCTGAGAGAGCTGTGGGG 1189  
Db 1292 AGGTCCCGCGGTGGGCTCTCTCATCTGCTGCTGCTGAGCTGAGAGAGCTGTGGGG 1351  
Qy 1190 AGAGGTGGCCCTTCAACCAAGCTCAAGGCTTGAAGCGCCCTTTGTGTTCAATC 1249  
Db 1352 AGAGGTGGCCCTTCAACCAAGCTCAAGGCTTGAAGCGCCCTTTGTGTTCAATC 1411  
Qy 1250 AGTCAACCCCTGAGGCGGTGCGGGCTGTCTCAATCTGGAGGAAGAGTGTGCTGA 1309  
Db 1412 AGTCAACCCCTGAGGCGGTGCGGGCTGTCTCAATCTGGAGGAAGAGTGTGCTGA 1471  
Qy 1310 AGCGCTTGTGCGGTGCTGTGAGAGAGTGTGAATAGAAATCAAGTGTGTTAGTGTG 1369  
Db 1472 AGCGCTTGTGCGGTGCTGTGAGAGAGTGTGAATAGAAATCAAGTGTGTTAGTGTG 1531  
Qy 1370 ACAGACTCATGTTTAAAGAGCTCGGGAGTCTGAGCTCCGTGCTTTCATGAACA 1429  
Db 1532 ACAGACTCATGTTTAAAGAGCTCGGGAGTCTGAGCTCCGTGCTTTCATGAACA 1591  
Qy 1430 TCCCTTCCCGAGACGCACTGTGAGAGTACTACCTTCTTCAACCTCAAGTCCGACG 1489  
Db 1592 TCCCTTCCCGAGACGCACTGTGAGAGTACTACCTTCTTCAACCTCAAGTCCGACG 1651  
Qy 1490 ACTTGTGTGCGCTTGTGTAATCACTGTGTTAGAGGGCTGAAATCGAGCAGCTCTG 1549  
Db 1652 ACTTGTGTGCGCTTGTGTAATCACTGTGTTAGAGGGCTGAAATCGAGCAGCTCTG 1711  
Qy 1550 CTCTGTAGTGAAGAACAAAGAGTTCATGAGACTTAAACAGGACGCTTCCATATCC 1609  
Db 1712 CTCTGTAGTGAAGAACAAAGAGTTCATGAGACTTAAACAGGACGCTTCCATATCC 1771

Qy 1610 ACTGCTTTGAGTGAAGGTTTCTTGTGTGGCTCGTGAAGGAGAGATGAAGAGCCAC 1669  
Db 1772 ACTGCTTTGAGTGAAGGTTTCTTGTGTGGCTCGTGAAGGAGAGATGAAGAGCCAC 1831  
Qy 1670 TGAAGTCTCTGAGGCTGTCCGCTCCCTGCGGGGTGAAGCAGAGCTTCTGACTGG 1729  
Db 1832 TGAAGTCTCTGAGGCTGTCCGCTCCCTGCGGGGTGAAGCAGAGCTTCTGACTGG 1891  
Qy 1730 TCTCTGTGTGGGTACAGGCTTAATGCACCACTCCAGAGAGACCTGAGCGCTTCC 1789  
Db 1892 TCTCTGTGTGGGTACAGGCTTAATGCACCACTCCAGAGAGACCTGAGCGCTTCC 1951  
Qy 1790 ACTGCTTTGAGACTCAAGCAAGAGTGTGCTGCTTGGCACTTAAACAGTTCACAG 1849  
Db 1952 ACTGCTTTGAGACTCAAGCAAGAGTGTGCTGCTTGGCACTTAAACAGTTCACAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAACCAAGAACTGGACTTGATAGCATCTTCTGTGCTCCAG 1909  
Db 2012 AAGTGTGCTTCCGATTAACCAAGAACTGGACTTGATAGCATCTTCTGTGCTCCAG 2071  
Qy 1910 ACTGTCCGTATTTGCGGAAAAATGGGTGATGTCAAAAGGATCTTCCAAAGATGAGT 1969  
Db 2072 ACTGTCCGTATTTGCGGAAAAATGGGTGATGTCAAAAGGATCTTCCAAAGATGAGT 2131  
Qy 1970 CCGCTGAGGCAATGTCTGTGGTCCCTCTATGATGATGCGGGAATGAACCTCAATTGAGAG 2029  
Db 2132 CCGCTGAGGCAATGTCTGTGGTCCCTCTATGATGATGCGGGAATGAACCTCAATTGAGAG 2191  
Qy 2030 AGTGGAGAAATTTCTGCTCCATGCTTGGCACTCCACCAACCTGCGGAGCTGACCTGG 2089  
Db 2192 AGTGGAGAAATTTCTGCTCCATGCTTGGCACTCCACCAACCTGCGGAGCTGACCTGG 2251  
Qy 2090 GCAGAGATTCCTGACAGAGCGGGCCATGAAGACCTGTGTGCGCAAGCTGAGGATCCA 2149  
Db 2252 GCAGAGATTCCTGACAGAGCGGGCCATGAAGACCTGTGTGCGCAAGCTGAGGATCCA 2311  
Qy 2150 CCTGCAAGATACAGACCTGATGTTTGAAGATGACAGATTAACCTCTGTGTGAGACAC 2209  
Db 2312 CCTGCAAGATACAGACCTGATGTTTGAAGATGACAGATTAACCTCTGTGTGAGACAC 2371  
Qy 2210 TCTGAGAAATGTCATGCGCAACCGTAACTTAAGATCTTCAACTTGGAGGCAACCA 2269  
Db 2372 TCTGAGAAATGTCATGCGCAACCGTAACTTAAGATCTTCAACTTGGAGGCAACCA 2431  
Qy 2270 TGAAGAAAGAGATGAAGATGCGGTGTAAGGCTTAAACACCCAAAATGTTGTGG 2329  
Db 2432 TGAAGAAAGAGATGAAGATGCGGTGTAAGGCTTAAACACCCAAAATGTTGTGG 2491  
Qy 2330 AGTCTTTGAGGCTGATTTGCTGTGATTAACCAATGCTGTTACTGAAGATCTCCAAA 2389  
Db 2492 AGTCTTTGAGGCTGATTTGCTGTGATTAACCAATGCTGTTACTGAAGATCTCCAAA 2551  
Qy 2390 TCTTTACGACTCCCGACGCTGAATCTGTAGCTGTGGCAGAGAACAGGTGAACAAC 2449  
Db 2552 TCTTTACGACTCCCGACGCTGAATCTGTAGCTGTGGCAGAGAACAGGTGAACAAC 2611  
Qy 2450 AGGAGTAATGCTCTGATGATGATGCTGAGAGTCTCCAGTGGCCCTGAGAAAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTGATGATGATGCTGAGAGTCTCCAGTGGCCCTGAGAAAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGAGATCAACAGCAGGTTTCCAGAGCTTGGCTCAAGCCTCTGCA 2569  
Db 2672 TACTGAGAGACTGTGAGATCAACAGCAGGTTTCCAGAGCTTGGCTCAAGCCTCTGCA 2731  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGTATTCACCAACAGCTGGGAAACGAGGTGTA 2629  
Db 2732 GCAACCGGAGCTTGAACAACCTGTGTATTCACCAACAGCTGGGAAACGAGGTGTA 2791  
Qy 2630 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTATCTGCAAGGCTGATGTGTAATC 2689  
Db 2792 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTATCTGCAAGGCTGATGTGTAATC 2851



QY 830 CTGCTCTCAAGATGACAAAGCTCTGCAAAAGCTGAGGCTGAGAGCAGCCTCGTTC A 889  
Db 1106 CTGCTCTCAAGATGACAAAGCTCTGCAAAAGCTGAGGCTGAGAGCAGCCTCGTTC A 1165  
QY 890 CCTCTCATACGAGCTGCTGAGAGAGGCTCTGCTCTCTGAGTCTTCTGATCTGTCACCG 949  
Db 1166 CCTCTCATACGAGCTGCTGAGAGAGGCTCTGCTCTCTGAGTCTTCTGATCTGTCACCG 1225  
QY 950 TCAGAGAGCTGGGACACAGAGAGCTCAGTCAGAGGCTGTCTCCCGGTTAACCTGTTAG 1009  
Db 1226 TCAGAGAGCTGGGACACAGAGAGGCTCAGTCAGAGGCTGTCTCCCGGTTAACCTGTTAG 1285  
QY 1010 TTAGAGGATCTCCGGGACAAAGAAATCACTTGCTCTTGAGCGCGGATGTGTAGC 1069  
Db 1286 TTAGAGGATCTCCGGGACAAAGAAATCACTTGCTCTTGAGCGCGGATGTGTAGC 1345  
QY 1070 ATCAGAGACACAGAGGTTGGTGGTCATGAAACAACCTGAGCTGCTCGACAGTGGC 1129  
Db 1346 ATCAGAGACACAGAGGTTGGTGGTCATGAAACAACCTGAGCTGCTCGACAGTGGC 1405  
QY 1130 AGGTGGCCGCGTGGGCTCTGTCATCTGCGTGGCCCTGACAGCTGACGAGCGTGGG 1189  
Db 1406 AGGTGGCCGCGTGGGCTCTGTCATCTGCGTGGCCCTGACAGCTGACGAGCGTGGG 1465  
QY 1190 AGAGCGTGGCCGCTTCAACCAAAAGCTCAGAGGCTGACGCGCTTTGTGTTCATC 1249  
Db 1466 AGAGCGTGGCCGCTTCAACCAAAAGCTCAGAGGCTGACGCGCTTTGTGTTCATC 1525  
QY 1250 AGCTCAACCTCTGAGAGGCTGCTCGGCGCTGTCTCAATCTGAGAGAAAGTGTCTGA 1309  
Db 1526 AGCTCAACCTCTGAGAGGCTGCTCGGCGCTGTCTCAATCTGAGAGAAAGTGTCTGA 1585  
QY 1310 AGCGCTTCTGCGCTATGCTGCTGAGAGGAGTGTGAATAGAGAGTCACTGTTTATGAGTG 1369  
Db 1586 AGCGCTTCTGCGCTATGCTGCTGAGAGGAGTGTGAATAGAGAGTCACTGTTTATGAGTG 1645  
QY 1370 ACGACCTCATGATTCAGAGAGCTCGGAGAGCTGAGAGCTCGTCTCTGTTTACATGAAACA 1429  
Db 1646 ACGACCTCATGATTCAGAGAGCTCGGAGAGCTGAGAGCTCGTCTCTGTTTACATGAAACA 1705  
QY 1430 TCTCTTCTCCAGACAGCCACTGTGTAGAGTACTACACTTCTTCCACTCACTCTCCAGG 1489  
Db 1706 TCTCTTCTCCAGACAGCCACTGTGTAGAGTACTACACTTCTTCCACTCACTCTCCAGG 1765  
QY 1490 ACTTCTGCGCGCTTGTACTACGTTGAGAGGAGCTGAGAAATGAGGCAAGCTCTGAGC 1549  
Db 1766 ACTTCTGCGCGCTTGTACTACGTTGAGAGGAGCTGAGAAATGAGGCAAGCTCTGAGC 1825  
QY 1550 CTCTGTACGTTGAGAGACAAAGAGGTCATGAGCTTAAACAGGCAAGCTTCCATATCC 1609  
Db 1826 CTCTGTACGTTGAGAGACAAAGAGGTCATGAGCTTAAACAGGCAAGCTTCCATATCC 1885  
QY 1610 ACTGCTTTGATGAGAGGCTTCTGTTTGGCTCTGTAGCGAAGAGTAAAGAGGCGAC 1669  
Db 1886 ACTGCTTTGATGAGAGGCTTCTGTTTGGCTCTGTAGCGAAGAGTAAAGAGGCGAC 1945  
QY 1670 TGAAGGTCCTGCTGGGCTGTCCCGTTGCCCTGGGGGTGAGAGAGGCTTGTGACATGGG 1729  
Db 1946 TGAAGGTCCTGCTGGGCTGTCCCGTTGCCCTGGGGGTGAGAGAGGCTTGTGACATGGG 2005  
QY 1730 TCTCTCTGTGGGTCAGAGGCTTAATGACCAACCCAGAGAGACACCTTGAACGCTTCC 1789  
Db 2006 TCTCTCTGTGGGTGAGAGGCTTAATGACCAACCCAGAGAGACACCTTGAACGCTTCC 2065  
QY 1790 ACTGCTTTTGAAGCTCAAGACAAAGAGTTGTTGCTTGGCATTTAAACAGCTTCCAG 1849  
Db 2066 ACTGCTTTTGAAGCTCAAGACAAAGAGTTGTTGCTTGGCATTTAAACAGCTTCCAG 2125  
QY 1850 AAGGTGGCTTCCGATTAACAGAGAACTTGAATGAGCATCTTCTTGGCTCCAGC 1909  
Db 2126 AAGGTGGCTTCCGATTAACAGAGAACTTGAATGAGCATCTTCTTGGCTCCAGC 2185  
QY 1910 ACTGTCGATTTTGGGAAAAATTCGGGTGAGATGTCAAGAGATCTTCCAAAGATGAGT 1969

Db 2186 ACTGTCGATTTTGGGAAAAATTCGGGTGAGATGTCAAGAGATCTTCCAAAGATGAGT 2245  
QY 1970 CCGCTGAGGATGTCTGTGGTCCCTCTATAGATGCGGATTAACAACCTCAATTAGAGAC 2029  
Db 2246 CCGCTGAGGATGTCTGTGGTCCCTCTATAGATGCGGATTAACAACCTCAATTAGAGAC 2305  
QY 2030 AGTGGAAAGATTTCTGCTCCATGCTTGGACCCACCAACCTGCGAGCTGAGACCTG 2089  
Db 2306 AGTGGAAAGATTTCTGCTCCATGCTTGGACCCACCAACCTGCGAGCTGAGACCTG 2365  
QY 2090 GCAGCAGATCTTGAACAGAGCGGACATGAAGACCTGTGTGCCAAGCTGAGCATCCCA 2149  
Db 2246 GCAGCAGATCTTGAACAGAGCGGACATGAAGACCTGTGTGCCAAGCTGAGCATCCCA 2425  
QY 2150 CCTGCAAGATACACACCTGTATGTTTGAATGACAGATTAACCCGTGTGTGAGACAC 2209  
Db 2426 CCTGCAAGATACACACCTGTATGTTTGAATGACAGATTAACCCGTGTGTGAGACAC 2485  
QY 2210 TCTGAGAAATGCTATGAGCCCAACCGTAACTTAAGATCCCTCACTTGGAGGACACCA 2269  
Db 2486 TCTGAGAAATGCTATGAGCCCAACCGTAACTTAAGATCCCTCACTTGGAGGACACCA 2545  
QY 2270 TGAAGAAAGAGATGTAAAGATGCGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2546 TGAAGAAAGAGATGTAAAGATGCGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2605  
QY 2330 AGTCTTTGAGGCTGGATTTGCTGTGATTTGACCCCATGCTGTACTGAAGATCTCCCAA 2389  
Db 2606 AGTCTTTGAGGCTGGATTTGCTGTGATTTGACCCCATGCTGTACTGAAGATCTCCCAA 2665  
QY 2390 TCTTTACGACCTCCCGACGCTGAAATCTCTGACCTTGACAGAAACAGAGTGAACAGC 2449  
Db 2666 TCTTTACGACCTCCCGACGCTGAAATCTCTGACCTTGACAGAAACAGAGTGAACAGC 2725  
QY 2450 AGGAGATATGCTCTGATGATGCTTTGAGATCTCCAGTGGCCCTGCAAGAGCTGA 2509  
Db 2726 AGGAGATATGCTCTGATGATGCTTTGAGATCTCCAGTGGCCCTGCAAGAGCTGA 2785  
QY 2510 TACTGAGAGACTGTGGATACACAGCCAGGTTTGCAGAGCTGAGCCCTGCTGCA 2559  
Db 2786 TACTGAGAGACTGTGGATACACAGCCAGGTTTGCAGAGCTGAGCCCTGCTGCA 2845  
QY 2570 GCAACCGGAGCTTGAACACCTGTGCTTATCCAACAAGCTGGGAAACGAAGTGTAA 2629  
Db 2846 GCAACCGGAGCTTGAACACCTGTGCTTATCCAACAAGCTGGGAAACGAAGTGTAA 2905  
QY 2630 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATC 2689  
Db 2906 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATC 2965  
QY 2690 AGTGCACCTGGAACACAGGCTGTGTGGTTTCTTGAACCTTGAGGTTAAGTCAAT 2749  
Db 2966 AGTGCACCTGGAACACAGGCTGTGTGGTTTCTTGAACCTTGAGGTTAAGTCAAT 3025  
QY 2750 GGTGAGGACCTTGAAGCTTATGATGAACCTGTGTGAAGACAATGAGGCTGAAGCTTGT 2809  
Db 3026 GGTGAGGACCTTGAAGCTTATGATGAACCTGTGTGAAGACAATGAGGCTGAAGCTTGT 3085  
QY 2810 GCGAGTTCATGAGAAACCATTTGTCAATCCAGACCTTGAAGTTGTTAAGTCTATC 2869  
Db 3086 GCGAGTTCATGAGAAACCATTTGTCAATCCAGACCTTGAAGTTGTTAAGTCTATC 3145  
QY 2870 TCAACCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACGACACCTGAAGA 2929  
Db 3146 TCAACCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACGACACCTGAAGA 3205  
QY 2930 GCTGTGATTCACAGGACAAATGCTCTGGGTGACGCTGGGTTGCTGCGCTGTGAGAGGAC 2989  
Db 3206 GCTGTGATTCACAGGACAAATGCTCTGGGTGACGCTGGGTTGCTGCGCTGTGAGAGGAC 3265  
QY 2990 TGAAGCAAAAGACAGTGTCTGACAGACTCGGATTGAAGGACATGTGACTGACTTCTG 3049

Db 3266 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGACTGTGACTGACTTCTG 3325  
Qy 3050 ATTTGCTGAGGCACTCTCTCTTGGCCCTTTCTTGCAACCGGATCTGACCAATCTTAAAC 3109  
Db 3336 ATTTGCTGAGGCACTCTCTCTTGGCCCTTTCTTGCAACCGGATCTGACCAATCTTAAAC 3385  
Qy 3110 TGGTCAGAAATTAATCTTCACTTCCCAAGAAATGATGAAAGTGTGCTTGGCTTGGCTGTC 3169  
Db 3386 TGGTCAGAAATTAATCTTCACTTCCCAAGAAATGATGAAAGTGTGCTTGGCTTGGCTGTC 3445  
Qy 3170 CCAGCTCTAATCTTCAAGATTAATTTGGGCTGTGAAAAATGGCAATACCTCTGTGCAATAGA 3229  
Db 3446 CCAGCTCTAATCTTCAAGATTAATTTGGGCTGTGAAAAATGGCAATACCTCTGTGCAATAGA 3505  
Qy 3230 AGCTCTGAGAGAAATGTCAGCTCAAGCCCGGAGTGTAAATTAAGAGTATTTGGCATT 3289  
Db 3506 AGCTCTGAGAGAAATGTCAGCTCAAGCCCGGAGTGTAAATTAAGAGTATTTGGCATT 3565  
Qy 3290 CTTTGTATGAAGATGACCG 3308  
Db 3566 CTTTGTATGAAGATGACCG 3584

RESULT 3  
AX704821 3926 bp DNA linear PAT 04-APR-2003  
LOCUS AX704821  
DEFINITION Sequence 1 from Patent EP1285964.  
ACCESSION AX704821  
VERSION AX704821.1 GI:29561487  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1  
AUTHORS Welts, B., Lesel, M., Peters-Kottig, M. and Beckmann, G.  
TITLE Human mater proteins  
JOURNAL Patent: EP 1285964-A 1 26-FEB-2003;  
SHERING AKTIBENSELSCHAFT (DE)  
FEATURES  
source 1..3926  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN

Query Match 72.3%; Score 2917; DB 6; Length 3926;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 650 ACCAGGGAATGTTCTCTACAGCTCTTCTCTCCCGTTAGAGAGATGACAGGGAAGAG 709  
Db 812 ACCAGGGAATGTTCTCTACAGCTCTTCTCTCCCGTTAGAGAGATGACAGGGAAGAG 871  
Qy 710 AGAGCAGTGTACAGAGTTCAATCTCAGAGAGTGGCCAGACTCCAGGCTCCGATGACG 769  
Db 872 AGAGCAGTGTACAGAGTTCAATCTCAGAGAGTGGCCAGACTCCAGGCTCCGATGACG 931  
Qy 770 AGATCATGTCCCGACAGAAAGCTGTGTTTATCATTTAGACGCTTTCATGACTGGCT 829  
Db 932 AGATCATGTCCCGACAGAAAGCTGTGTTTATCATTTAGACGCTTTCATGACTGGCT 991  
Qy 830 CTGTCTCAAAATGACAAAGCTGTGCAAAAGCTGGGCTGAGAGAGAGCTCCGTTCA 889  
Db 992 CTGTCTCAAAATGACAAAGCTGTGCAAAAGCTGGGCTGAGAGAGAGCTCCGTTCA 1051  
Qy 890 CCTCATACGAGCTGTGCTGAGAAAGTCTGCTCCCTGAGTCTTCTGATTCGACCG 949  
Db 1052 CCTCATACGAGCTGTGCTGAGAAAGTCTGCTCCCTGAGTCTTCTGATTCGACCG 1111  
Qy 950 TCAGAGAGTGGGACAGAGAACTCAAGTCAAGAGTGTGTCTCCGTTAATCTGTTAG 1009  
Db 1112 TCAGAGAGTGGGACAGAGAACTCAAGTCAAGAGTGTGTCTCCGTTAATCTGTTAG 1171  
Qy 1010 TTAGAGGAATCTCCGAGGGAACAAAGATCCACTTCTCTTGAAGCGGAGTTGTGAGC 1065  
Db 1172 TTAGAGGAATCTCCGAGGGAACAAAGATCCACTTCTCTTGAAGCGGAGTTGTGAGC 1231  
Qy 1070 ATCAGAGACAAGAGGTTGGTGGATCATGAACAACGAGTGTCTGACCAAGTGC 1129  
Db 1232 ATCAGAGACAAGAGGTTGGTGGATCATGAACAACGAGTGTCTGACCAAGTGC 1291  
Qy 1130 AGGTCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGCGTGGGG 1189  
Db 1292 AGGTCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGCGTGGGG 1351  
Qy 1190 AGAGGTGGCCCGCTTCAACCAACAGCTCAAGGCTGCAAGCGGCTTTGTGTTATC 1249  
Db 1352 AGAGGTGGCCCGCTTCAACCAACAGCTCAAGGCTGCAAGCGGCTTTGTGTTATC 1411  
Qy 1250 AGCTCACCCCTGAGAGGCTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTGGCTGCA 1309  
Db 1412 AGCTCACCCCTGAGAGGCTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTGGCTGCA 1471  
Qy 1310 AGCGCTTCCGCTATGCTGTGAGAGGAGTGTGAATGAGAGTCACTGTTTATGAGT 1369  
Db 1472 AGCGCTTCCGCTATGCTGTGAGAGGAGTGTGAATGAGAGTCACTGTTTATGAGT 1531  
Qy 1370 ACGACTCATGTTCAAGGACTCGGGAGTCTGAGCTCGTGTCTGTTTCAATGAACA 1429  
Db 1532 ACGACTCATGTTCAAGGACTCGGGAGTCTGAGCTCGTGTCTGTTTCAATGAACA 1591  
Qy 1430 TCTTCTCCAGACAGCCACTGTGAGAGTCTACCTTCTTCACTCACTCAAG 1489  
Db 1592 TCTTCTCCAGACAGCCACTGTGAGAGTCTACCTTCTTCACTCACTCAAG 1651  
Qy 1490 ACTTCTGCGGCTTGTACTACGTGTTAGAGGCTGGAATGAGCACTCTCTGCC 1549  
Db 1652 ACTTCTGCGGCTTGTACTACGTGTTAGAGGCTGGAATGAGCACTCTCTGCC 1711  
Qy 1550 CTCTGTAGTTGAGAGCAAAAGTGTATGAGCTTTAAACAGGCAAGCTTCAATATCC 1609  
Db 1712 CTCTGTAGTTGAGAGCAAAAGTGTATGAGCTTTAAACAGGCAAGCTTCAATATCC 1771  
Qy 1670 ACTGCTTTGAGATGAAGGTTTCTTGTGGCTTCGAGGAGAGCTAAGAGGCGAC 1669  
Db 1772 ACTGCTTTGAGATGAAGGTTTCTTGTGGCTTCGAGGAGAGCTAAGAGGCGAC 1831  
Qy 1670 TGAAGTCTGCTGAGCTGTCCGTTCCCTGAGGAGTGAAGCAAGCTTCTGACTGG 1729  
Db 1832 TGAAGTCTGCTGAGCTGTCCGTTCCCTGAGGAGTGAAGCAAGCTTCTGACTGG 1891



Qy 1730 TCTCTGTGTGAGTCAAGGCTTAATGACACCCAGAGAGACCTTGACGCTTCC 1789  
Db 1692 TCTCTGTGTGAGTCAAGGCTTAATGACACCCAGAGAGACCTTGACGCTTCC 1951  
Qy 1790 ACTGTCTTTTGAGACTCAAGCAAAAGTTTGTTCGCTTGACATTAACGCTTCAAG 1849  
Db 1952 ACTGTCTTTTGAGACTCAAGCAAAAGTTTGTTCGCTTGACATTAACGCTTCAAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAAACAGAACTGATAGCATCTTCTTGCTCCAGC 1909  
Db 2012 AAGTGTGCTTCCGATTAAACAGAACTGATAGCATCTTCTTGCTCCAGC 2071  
Qy 1910 ACTGTCTTTTGAGAAAAATTCGGGTGATGTCAAAAGGATCTTCCAAAGATAGT 1969  
Db 2072 ACTGTCTTTTGAGAAAAATTCGGGTGATGTCAAAAGGATCTTCCAAAGATAGT 2131  
Qy 1970 CCGCTGAGGCAATGCTGTGTGCTCTCTATGATGCGGATTAAGACCTCATTTGAGAGC 2029  
Db 2132 CCGCTGAGGCAATGCTGTGTGCTCTCTATGATGCGGATTAAGACCTCATTTGAGAGC 2191  
Qy 2030 AAGTGGAAAGATTTTGTGCTCCATGCTTGGACCCACCTGCGGAGCTGGACCTGG 2089  
Db 2192 AAGTGGAAAGATTTTGTGCTCCATGCTTGGACCCACCTGCGGAGCTGGACCTGG 2251  
Qy 2090 GCAGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAAGCATCCA 2149  
Db 2252 GCAGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAAGCATCCA 2311  
Qy 2150 CCTGTCAAGATACAGACCTGTGATTTAGAAATGCAAGATTAACCTGTGTGAGACAC 2209  
Db 2312 CCTGTCAAGATACAGACCTGTGATTTAGAAATGCAAGATTAACCTGTGTGAGACAC 2371  
Qy 2210 TCTGAGAAATGTCATGAGCAACCGTAACTTAAGTCCCTCAACTTGGAGAGACCCAGC 2269  
Db 2372 TCTGAGAAATGTCATGAGCAACCGTAACTTAAGTCCCTCAACTTGGAGAGACCCAGC 2431  
Qy 2270 TGAAGAAAGAGATGTAAGATGCGGTGTAAGGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2432 TGAAGAAAGAGATGTAAGATGCGGTGTAAGGCTTAAACACCCAAATGTTTGTGG 2491  
Qy 2330 AGTCTTTGAGGCTGTGATGCTGTGATGACCCATGCTGTTAACTGTAAGATCTCCCAA 2389  
Db 2492 AGTCTTTGAGGCTGTGATGCTGTGATGACCCATGCTGTTAACTGTAAGATCTCCCAA 2551  
Qy 2390 TCTTACGACCTCCCGCAGCGTGAATCTGAGCCTGCGAGAAACAAGGTGACAGC 2449  
Db 2552 TCTTACGACCTCCCGCAGCGTGAATCTGAGCCTGCGAGAAACAAGGTGACAGC 2611  
Qy 2450 AGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTCTGAGAGAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTCTGAGAGAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGAGCATCAAGCAGCGGTTGCGAAGTCTGAGCCTTCCCTCA 2569  
Db 2672 TACTGAGAGACTGTGAGCATCAAGCAGCGGTTGCGAAGTCTGAGCCTTCCCTCA 2731  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAAGGCTGGGAAACAAGGTGAA 2629  
Db 2732 GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAAGGCTGGGAAACAAGGTGAA 2791  
Qy 2630 ATCTACTGTGTGATCATTAGAGCTTCCCACTGTAGTCTGAGAGGCTGATGCTGAATC 2689  
Db 2792 ATCTACTGTGTGATCATTAGAGCTTCCCACTGTAGTCTGAGAGGCTGATGCTGAATC 2851  
Qy 2690 AGTGTCACTTGAACAAGGCTGTGCTGTTTCTTGAACCTTGCCTTATGGGTAATCAT 2749  
Db 2852 AGTGTCACTTGAACAAGGCTGTGCTGTTTCTTGAACCTTGCCTTATGGGTAATCAT 2911  
Qy 2750 GAGTGAAGCAGCTGAGCTTATGAGCAACCTGTGGAAGACAATGGAGCTGAGCTTGT 2809  
Db 2912 GAGTGAAGCAGCTGAGCTTATGAGCAACCTGTGGAAGACAATGGAGCTGAGCTTGT 2971  
Qy 2810 GCGAGTCAATGAGAAACCATCTTGTCACTTCAAGACCTGAGATTGGTAAAGTGCATC 2869

Db 2972 GCGAGTCAATGAGAAACCATCTTGTCACTTCCAGAGACTGAGATTGTAAGTGCATC 3031  
Qy 2870 TCACCGCGCGTGTGTGAGAGTGTCTGCTGTGATCTGAGAGAGACACCTGAAGA 2929  
Db 3032 TCACCGCGCGTGTGTGAGAGTGTCTGCTGTGATCTGAGAGAGACACCTGAAGA 3091  
Qy 2930 GCTGTGATCTCAACGAAATGCCCTGGGTGACAGGTGGGTTGTGCTGCTGCGAGGAC 2989  
Db 3092 GCTGTGATCTCAACGAAATGCCCTGGGTGACAGGTGGGTTGTGCTGCTGCGAGGAC 3151  
Qy 2990 TGAAGCAAAAGACAGTGTCTGACAGACTCGGCTTGAAGGATGTGACTGACTTCTG 3049  
Db 3152 TGAAGCAAAAGACAGTGTCTGACAGACTCGGCTTGAAGGATGTGACTGACTTCTG 3211  
Qy 3050 ATTGTGTGAGGCACTCTCCCTGGCCCTTTCCTGCAACCGGCAATGACAGCTAAAC 3109  
Db 3212 ATTGTGTGAGGCACTCTCCCTGGCCCTTTCCTGCAACCGGCAATGACAGCTAAAC 3271  
Qy 3110 TGTGTCAAGAAATTAACCTTCACTCCAAAGAAATGAAGTGTGTTCCGCTTGTGCTGTC 3169  
Db 3272 TGTGTCAAGAAATTAACCTTCACTCCAAAGAAATGAAGTGTGTTCCGCTTGTGCTGTC 3331  
Qy 3170 CCAGCTTAACCTTACAGATTAATGGGCTGTGAAATGGCAGTACCTGTGCAATTAAGA 3229  
Db 3332 CCAGCTTAACCTTACAGATTAATGGGCTGTGAAATGGCAGTACCTGTGCAATTAAGA 3391  
Qy 3230 AGCTGTGAGGAAAGTCACTCAAGCCCGAGTGGTAATGAGGATGTTGGGATT 3289  
Db 3392 AGCTGTGAGGAAAGTCACTCAAGCCCGAGTGGTAATGAGGATGTTGGGATT 3451  
Qy 3290 CTTTGTGATGAAGATGACCG 3308  
Db 3452 CTTTGTGATGAAGATGACCG 3470

RESULT 4  
AX704823 3830 bp DNA linear PAT 04-APR-2003  
LOCUS AX704823  
DEFINITION Sequence 3 from Patent EP1285964.  
ACCESSION AX704823  
VERSION AX704823.1 GI:29561488  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Weiss, B., Lessl, M., Peters-Kotzig, M. and Beckmann, G.  
AUTHORS Human mater proteins  
TITLE Patent: EP 1285964-A 3 26-FEB-2003;  
JOURNAL SCHERING AKTENSELSCHAFT (DE)  
FEATURES  
Location/Qualifiers  
1. .3830  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 70.9%; Score 2862; DB 6; Length 3830;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2962; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 345 AGAATTTTCAACACTATGGAACAAGAGTGCACAGCAGCAGAGACAGAAAGG 404  
Db 450 AGAATTTTCAACACTATGGAACAAGAGTGCACAGCAGCAGAGACAGAAAGG 509  
Qy 405 ACATGAGGTGACACATGAGGACTCAAGAGTCAAGTATGACCAATTCGTGAGAGGA 464  
Db 510 ACATGAGGTGACACATGAGGACTCAAGAGTCAAGTATGACCAATTCGTGAGAGGA 569  
Qy 465 GGATGTAGTGTGATGTTTGAAGAACTGCTGCTGACTGGCCGGAATGCAACGTTGGC 524



Db 570 GGATGACGTCGATGTTTAAAAACAGCTGCTGACATGCGCGGAAATGCAACGTTGCG 629  
Qy 525 TGGTCTTTTATTCAGACCGGTGGGGCTTCGGGCTCCGACGGTGTCTGACGGAA 584  
Db 630 TGGTCTTTTATTCAGACCGGTGGGGCTTCGGGCTCCGACGGTGTCTGACGGAA 689  
Qy 585 GTGAGAAATTGGGAAATCGGCTCTAGCGAGAAAGATCGTGTGCTGGGCGCAAGGTG 644  
Db 690 GTGAGAAATTGGGAAATCGGCTCTAGCGAGAAAGATCGTGTGCTGGGCGCAAGGTG 749  
Qy 645 ACTCTACAGGGAATGTTCTCTACGCTCTTCTCTCTCCCGTTAGAGATGACGGAA 704  
Db 750 ACTCTACAGGGAATGTTCTCTCTAGGCTCTTCTCTCTCCCGTTAGAGATGACGGAA 809  
Qy 705 GAAAGAGACATGTCTACAGATTCCTCTCAGGAGTGGCCAGACTCCAGGCTCCGT 764  
Db 810 GAAAGAGACATGTCTACAGATTCCTCTCAGGAGTGGCCAGACTCCAGGCTCCGT 869  
Qy 765 GACGAGATCATGTCCCGACAGAAAGGCTGTGTTCACTTGAACGTTTGATGACCT 824  
Db 870 GACGAGATCATGTCCCGACAGAAAGGCTGTGTTCACTTGAACGTTTGATGACCT 929  
Qy 825 GGGCTCTGTCTCAACATGACAAAGCTCTGCAAGATGCGGCTGAGAACAGCTCC 884  
Db 930 GGGCTCTGTCTCAACATGACAAAGCTCTGCAAGATGCGGCTGAGAACAGCTCC 989  
Qy 885 GTTCAACCTCATACGCAATGTCTGTGAGAGAGTCTGCTCTCTGATCTTCTGATCT 944  
Db 990 GTTCAACCTCATACGCAATGTCTGTGAGAGAGTCTGCTCTCTGATCTTCTGATCT 1049  
Qy 945 CACCGTCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACT 1004  
Db 1050 CACCGTCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACT 1109  
Qy 1005 GTTAGTAGAGGATCTCCGGGAAACAAAGATCCACTTCTCTTGAAGCGGATG 1064  
Db 1110 GTTAGTAGAGGATCTCCGGGAAACAAAGATCCACTTCTCTTGAAGCGGATG 1169  
Qy 1065 TGAGCATGAAAGACAAAGGTTGCTGATCATGAACAAACGTTGACTCTGACCA 1124  
Db 1170 TGAGCATGAAAGACAAAGGTTGCTGATCATGAACAAACGTTGACTCTGACCA 1229  
Qy 1125 GTGCAAGGTCGGGCTGCTCTCATCTGCGTGGCCCTGACGCTGACAGAGTGT 1184  
Db 1230 GTGCAAGGTCGGGCTGCTCTCATCTGCGTGGCCCTGACGCTGACAGAGTGT 1289  
Qy 1185 GGGGAGAGCGTGGGCTCTTCAACAAACGCTCAAGGCTGACGCGCTTTGTGT 1244  
Db 1290 GGGGAGAGCGTGGGCTCTTCAACAAACGCTCAAGGCTGACGCGCTTTGTGT 1349  
Qy 1245 TCATCAGCTCACCCCTCGAGGCTGTGCTCGGCTGTCTCAATCTGAGAAAGATTGT 1304  
Db 1350 TCATCAGCTCACCCCTCGAGGCTGTGCTCGGCTGTCTCAATCTGAGAAAGATTGT 1409  
Qy 1305 CCTGAAGGCTTCTGCGGCTGATGCTGTGAGAGGATGAGAAATGAGTGTGGA 1364  
Db 1410 CCTGAAGGCTTCTGCGGCTGATGCTGTGAGAGGATGAGAAATGAGTGTGGA 1469  
Qy 1365 TGGTACGACATCATGTTCAAGGATCTCGGAGATCTGAGCTCCGCTCTGTTCACAT 1424  
Db 1470 TGGTACGACATCATGTTCAAGGATCTCGGAGATCTGAGCTCCGCTCTGTTCACAT 1529  
Qy 1425 GAAACATCTTCTCCAGACAGCTGTGAGAGATCAACCTTCTTCAACCTGACT 1484  
Db 1530 GAAACATCTTCTCCAGACAGCTGTGAGAGATCAACCTTCTTCAACCTGACT 1589  
Qy 1485 CCAGACATCTGCTGCGGCTTGTACTAGTGTTAAGAGGCTGAGAAATGAGGCTCT 1544  
Db 1590 CCAGACATCTGCTGCGGCTTGTACTAGTGTTAAGAGGCTGAGAAATGAGGCTCT 1649  
Qy 1545 CTGCGCTCTGTAGTTGAGAGACAAAGAGTCCATGAGCTTTAAACAGGCGCTTCA 1604  
Db 1650 CTGCGCTCTGTAGTTGAGAGACAAAGAGTCCATGAGCTTTAAACAGGCGCTTCA 1709

Qy 1605 TATCCACTCGTTTNGATGAAGCGTTTCTTGTGTTGGCTCGTGAACGAAGATGAAGAG 1664  
Db 1710 TATCCACTCGTTTNGATGAAGCGTTTCTTGTGTTGGCTCGTGAACGAAGATGAAGAG 1769  
Qy 1665 GCCACTGAGAGTCTGCTGAGGCTGTCCGCTTCCCTGGGGGGTGAAGCAGAACTTCTCA 1724  
Db 1770 GCCACTGAGAGTCTGCTGAGGCTGTCCGCTTCCCTGGGGGGTGAAGCAGAACTTCTCA 1829  
Qy 1725 CTGGTCTCTCTGTTGGGTCAAGACCTTAATGCCAACCCAGAGACACCTTGAGCGC 1784  
Db 1830 CTGGTCTCTCTGTTGGGTCAAGACCTTAATGCCAACCCAGAGACACCTTGAGCGC 1889  
Qy 1785 CTTCACATCTCTTTTCAAGACTCAAGACAAAGATTGTGCTTGGCTTTAAACGCTT 1844  
Db 1890 CTTCACATCTCTTTTCAAGACTCAAGACAAAGATTGTGCTTGGCTTTAAACGCTT 1949  
Qy 1845 CCAAGAGTGGGCTCTCGATTAACAGAACCTGACTTGAATGATCTTCTCTGCT 1904  
Db 1950 CCAAGAGTGGGCTCTCGATTAACAGAACCTGACTTGAATGATCTTCTCTGCT 2009  
Qy 1905 CCAGCACTGTCCGTAATTTGCGGAAATTCGGGTGATGTCAAGGATCTTCCAAAGAG 1964  
Db 2010 CCAGCACTGTCCGTAATTTGCGGAAATTCGGGTGATGTCAAGGATCTTCCAAAGAG 2069  
Qy 1965 TGAATCCGCTGAGGATGTCTGTGTGCTCTCTATGATGCGGATTAAGACCTCATTTGA 2024  
Db 2070 TGAATCCGCTGAGGATGTCTGTGTGCTCTCTATGATGCGGATTAAGACCTCATTTGA 2129  
Qy 2025 GAGAGAGTGGGAAATTTCTGCTCAATGCTTGGACCCACCAACCTGCGGAGCTGGA 2084  
Db 2130 GAGAGAGTGGGAAATTTCTGCTCAATGCTTGGACCCACCAACCTGCGGAGCTGGA 2189  
Qy 2085 CTTGGGAGAGCATCTCTGACAGAGCGGACCATGAAGACCTGTGTGCAAGCTGAGCA 2144  
Db 2190 CTTGGGAGAGCATCTCTGACAGAGCGGACCATGAAGACCTGTGTGCAAGCTGAGCA 2249  
Qy 2145 TCCACCTGCAAGTACAGACCTGTGATTTGAAGATGACAGATTACCTGTGTGCA 2204  
Db 2250 TCCACCTGCAAGTACAGACCTGTGATTTGAAGATGACAGATTACCTGTGTGCA 2309  
Qy 2205 GCACTCTGAGAGTCTGATGAGGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGAC 2264  
Db 2310 GCACTCTGAGAGTCTGATGAGGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGAC 2369  
Qy 2265 CCACCTGAAGAGATGTAAGATGAGCTGTGAAAGCTTTAAACACCCAAATGTTT 2324  
Db 2370 CCACCTGAAGAGATGTAAGATGAGCTGTGAAAGCTTTAAACACCCAAATGTTT 2429  
Qy 2325 GTTGAAGTCTTTGAGGCTGATGCTGTGATGATGATGATGATGATGATGATGATG 2384  
Db 2430 GTTGAAGTCTTTGAGGCTGATGCTGTGATGATGATGATGATGATGATGATGATG 2489  
Qy 2385 CCAATCTTTACGACCTTCCCGACGCTGAATCTCTGAGCTTGGCAAGAAACAGGTGAC 2444  
Db 2490 CCAATCTTTACGACCTTCCCGACGCTGAATCTCTGAGCTTGGCAAGAAACAGGTGAC 2549  
Qy 2445 AGACCAAGGATTAATGCTCTCACTGATGCTTGAAGATCTCCAGTGGCGCTGCAAG 2504  
Db 2550 AGACCAAGGATTAATGCTCTCACTGATGCTTGAAGATCTCCAGTGGCGCTGCAAG 2609  
Qy 2505 GCTGATCTGAGAGATGTTGGCATCAAGCCAGGGTTGCCAGAGTCTGGCTTCAAGCT 2564  
Db 2610 GCTGATCTGAGAGATGTTGGCATCAAGCCAGGGTTGCCAGAGTCTGGCTTCAAGCT 2669  
Qy 2565 GCTGAGCAACCGGAGCTTGAACACCTGTGCTTATCAACAAACGCTTGGGAAACGAGG 2624  
Db 2670 GCTGAGCAACCGGAGCTTGAACACCTGTGCTTATCAACAAACGCTTGGGAAACGAGG 2729  
Qy 2625 TGTAAATCTACTGTGATCATGAGGCTTCCCACTGTAGTCTGAGAGGCTGATGCT 2684  
Db 2730 TGTAAATCTACTGTGATCATGAGGCTTCCCACTGTAGTCTGAGAGGCTGATGCT 2789

QY 2685 GAATCAGTCCACCTGACACGCGTGTGTTTCTTGCACTTGGCTTATGGGTAA 2744  
DB 2790 GAATCAGTCCACCTGACACGCGTGTGTTTCTTGCACTTGGCTTATGGGTAA 2849  
QY 2745 CTCTATGCTGACGCACTGAGCTTATGATGAACTTGTGAAAGCAATGGCGTGAAGCT 2804  
DB 2850 CTCTATGCTGACGCACTGAGCTTATGATGAACTTGTGAAAGCAATGGCGTGAAGCT 2909  
QY 2805 TCTGTGGAGGTCATGAGAGAACCATCTGTATCTCAGAGCCGTGGATGGTAAATG 2864  
DB 2910 TCTGTGGAGGTCATGAGAGAACCATCTGTATCTCAGAGCCGTGGATGGTAAATG 2969  
QY 2865 TCATCTCACCGCCGCTGTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACACT 2924  
DB 2970 TCATCTCACCGCCGCTGTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACACT 3029  
QY 2925 GAAAGAGCTGATCTCAGGACATGCTTGGGTGAAGGTGGGTGCTGCGCTGTGGA 2984  
DB 3030 GAAAGAGCTGATCTCAGGACATGCTTGGGTGAAGGTGGGTGCTGCGCTGTGGA 3089  
QY 2985 GGAAGCTGAAAGAAAGAACAGTGTCTGACGAGACTGCGGTTGAAGGATGAGACTGAC 3044  
DB 3090 GGAAGCTGAAAGAAAGAACAGTGTCTGACGAGACTGCGGTTGAAGGATGAGACTGAC 3149  
QY 3045 TTCTGATTTCTGTGAGGCACTCTCCCTTGGCCCTTCTGCAACCGGCACTGACCACT 3104  
DB 3150 TTCTGATTTCTGTGAGGCACTCTCCCTTGGCCCTTCTGCAACCGGCACTGACCACT 3209  
QY 3105 AAACCTGGTGCAGATTAATTCTGATCCCAAGAAATGATGAAGTGTGTGCGCTTTGC 3164  
DB 3210 AAACCTGGTGCAGATTAATTCTGATCCCAAGAAATGATGAAGTGTGTGCGCTTTGC 3269  
QY 3165 CTGTCCCAAGCTCTAATTCAAGATTAATTGGGCTGTGAAATGGAGTACCCTGTGCAAT 3224  
DB 3270 CTGTCCCAAGCTCTAATTCAAGATTAATTGGGCTGTGAAATGGAGTACCCTGTGCAAT 3329  
QY 3225 AAGGAGCTGCTGAGGAAAGTGCAGTACTCAAGCCCGAGTCTGTAAATGACGATGTTG 3284  
DB 3330 AAGGAGCTGCTGAGGAAAGTGCAGTACTCAAGCCCGAGTCTGTAAATGACGATGTTG 3389  
QY 3285 GCATTCTTTTGTGATGAGTACCG 3308  
DB 3390 GCATTCTTTTGTGATGAGTACCG 3413

RESULT 5  
AX459881  
LOCUS AX459881 5859 bp DNA linear PAT 30-NOV-2002  
DEFINITION Sequence 24 from Patent WO240668.  
ACCESSION AX459881  
VERSION AX459881.1 GI:21725649  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1  
AUTHORS Techopp, J. and Martinon, F.  
TITLE Proteins and dna sequences underlying these proteins used for  
JOURNAL treating inflammation  
Patent: WO 0240668-A 24 23-MAY-2002;  
Apotech Research and Development Ltd. (CH)  
FEATURES  
source location/Qualifiers  
1. 5859  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="NALP8/Pyl2--Pyl8\_dna"

ORIGIN  
Query Match 69.5%; Score 2803; DB 6; Length 5859;  
Best Local Similarity .99.9%; Pred. No. 0;  
Matches 2903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 290 AAATTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGACAGAAACAAGAAA 349  
DB 2287 AAATTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGACAGAAACAAGAAA 2446  
QY 350 TTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGACAGAAACAAGAACATG 409  
DB 2447 TTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGACAGAAACAAGAACATG 2506  
QY 410 GAGGTGACACATGGGACTTACAGAGTCAAGTGAACCAATTGCTGAGAGGAGATG 469  
DB 2507 GAGGTGACACATGGGACTTACAGAGTCAAGTGAACCAATTGCTGAGAGGAGATG 2566  
QY 470 TACGTCGTAGTTTGAAGAAACACTGCTGTGATGAGCCGGAATGCAACGTTGCTGTG 529  
DB 2567 TACGTCGTAGTTTGAAGAAACACTGCTGTGATGAGCCGGAATGCAACGTTGCTGTG 2626  
QY 530 CTTTGTGATTCAGACCGGTGGGGCTTCCGAGCTCGACAGTGTGTTCTGACGGAAGTCA 589  
DB 2627 CTTTGTGATTCAGACCGGTGGGGCTTCCGAGCTCGACAGTGTGTTCTGACGGAAGTCA 2686  
QY 590 GAATTTGGGAAATCGGCTCTAGCCAGAGATCGTGTGCTGGGGCCAGGTGGACTCT 649  
DB 2687 GAATTTGGGAAATCGGCTCTAGCCAGAGATCGTGTGCTGGGGCCAGGTGGACTCT 2746  
QY 650 ACCAGGAAATGTTCTCTAGCTCTTCTTCCCTCCGTTAGAGAGATCAGCGAAGAG 709  
DB 2747 ACCAGGAAATGTTCTCTAGCTCTTCTTCCCTCCGTTAGAGAGATCAGCGAAGAG 2806  
QY 710 AGACAGGTGTCAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGAGTCCGGTGAACG 769  
DB 2807 AGACAGGTGTCAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGAGTCCGGTGAACG 2866  
QY 770 AGATCATGTCCCGACAGAGAAAGCTGTGTTCATCATTTGACGTTTGATGACCTGGGCT 829  
DB 2867 AGATCATGTCCCGACAGAGAAAGCTGTGTTCATCATTTGATGACCTGGGCT 2926  
QY 830 CTGTCTCAACAATGACACAAAGCTCTGCAAGAGTGGGCTGAGAAACAGCTCTGTTCA 889  
DB 2927 CTGTCTCAACAATGACACAAAGCTCTGCAAGAGTGGGCTGAGAAACAGCTCTGTTCA 2986  
QY 890 CCTCATACGAGAGTCTGTCTGAGAGAGTCTCTCTGAGTCTTCTGATGCTCAACG 949  
DB 2987 CCTCATACGAGAGTCTGTCTGAGAGAGTCTCTCTGAGTCTTCTGATGCTCAACG 3046  
QY 950 TCAGAGACGTGGGACAGAGAGTCAAGTCAAGAGTGTGTGCTCCCGTTACCTGTAG 1009  
DB 3047 TCAGAGACGTGGGACAGAGAGTCAAGTCAAGAGTGTGTGCTCCCGTTACCTGTAG 3106  
QY 1010 TTAGAGAAATCTCCGGGGAACAAAGAAATCACTTGTCTCTTGAAGCGGGAATTTGTGAGC 1069  
DB 3107 TTAGAGAAATCTCCGGGGAACAAAGAAATCACTTGTCTCTTGAAGCGGGAATTTGTGAGC 3166  
QY 1070 ATCAGAAAGACAAAGGTTGGTGTGATGAAACAACCTGAGTCTGTGACCAAGTGGC 1129  
DB 3167 ATCAGAAAGACAAAGGTTGGTGTGATGAAACAACCTGAGTCTGTGACCAAGTGGC 3226  
QY 1130 AGGTGCGCGCGGTGGGCTCTCATCTGCGGTGGCCGAGCTGCAAGAGAGTGTGGGGG 1189  
DB 3227 AGGTGCGCGCGGTGGGCTCTCATCTGCGGTGGCCGAGCTGCAAGAGAGTGTGGGGG 3286  
QY 1190 AGAGCGTGGCGCCCTTCAACCAAGCTCAAGGCTGCAAGCGCTTTTGTGTTATC 1249  
DB 3287 AGAGCGTGGCGCCCTTCAACCAAGCTCAAGGCTGCAAGCGCTTTTGTGTTATC 3346  
QY 1250 AGCTACCCCTCTGAGGCGTGTGCTGCTCAATCTGAGAGAAAGTGTCTCTGA 1309  
DB 3347 AGCTACCCCTCTGAGGCGTGTGCTGCTCAATCTGAGAGAAAGTGTCTCTGA 3406  
QY 1310 AGCGCTTGTGCGGATGAGTGTGAGAGAGTGTGAAATAGAGATCAGTGTGATGATG 1369  
DB 3407 AGCGCTTGTGCGGATGAGTGTGAGAGAGTGTGAAATAGAGATCAGTGTGATGATG 3466



ORIGIN

/db\_xref="taxon:9606"  
/note="NALP5/Py8.dna /GENSCAN\_predicted\_CDS\_1/ 6939\_bp"

Query Match 69.5%; Score 2803; DB 6; Length 6939;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 280 AAATTTCAGGCTATGGAACAAGAGGTGCGACAGGACGACAGACAGAAACAAGAAA 349  
DB 3467 AAATTTCAGGCTATGGAACAAGAGGTGCGACAGGACGACAGACAGAAACAAGAAA 3526  
QY 350 TTTTCAAGCTATGGAACAAGAGGTGCGACAGGACGACAGACAGAAACAAGACATG 409  
DB 3527 TTTTCAAGCTATGGAACAAGAGGTGCGACAGGACGACAGACAGAAACAAGACATG 3586  
QY 410 GAGGTGACATGAGGACTTCAAGAGTCACTGATGACCAAAATTCCCTGAGAGAGAGATG 469  
DB 3587 GAGGTGACATGAGGACTTCAAGAGTCACTGATGACCAAAATTCCCTGAGAGAGAGATG 3646  
QY 470 TACGTCGATGTTTGAANAACCTGCTGACAGTGGCCGGAATGCAAACTTGGCTGGTG 529  
DB 3647 TACGTCGATGTTTGAANAACCTGCTGACAGTGGCCGGAATGCAAACTTGGCTGGTG 3706  
QY 530 CTTTGTATTCAGACCGGTGGGCTTCCGACCTCGACAGGTGGTTCGACGGAAGTCAG 589  
DB 3707 CTTTGTATTCAGACCGGTGGGCTTCCGACCTCGACAGGTGGTTCGACGGAAGTCAG 3766  
QY 590 GAATTGGGAATCGGCTCTAGCCAGAGAGATGTGTCTGTGCTGGGCGCAAGTGGACTCT 649  
DB 3767 GAATTGGGAATCGGCTCTAGCCAGAGAGATGTGTCTGTGCTGGGCGCAAGTGGACTCT 3826  
QY 650 ACCAGGAGATGTTCTCTACGCTCTTCTCCCTCCGTTAGAGAGATGACGAGGAAAGAG 709  
DB 3827 ACCAGGAGATGTTCTCTACGCTCTTCTCCCTCCGTTAGAGAGATGACGAGGAAAGAG 3886  
QY 710 AGAGCAGTGTCAAGAGTTCATCTCCAGAGATGAGCCAGACTCCAGGCTCCGATGACG 769  
DB 3887 AGAGCAGTGTCAAGAGTTCATCTCCAGAGATGAGCCAGACTCCAGGCTCCGATGACG 3946  
QY 770 AGATCATGTCCCGACCAAGAAAGCTGTGTTCATTCATTCAGCGTTTCATTCAGCTGGCT 829  
DB 3947 AGATCATGTCCCGACCAAGAAAGCTGTGTTCATTCATTCAGCGTTTCATTCAGCTGGCT 4006  
QY 830 CTGTCTCAACAATGACCAAAAGCTCTGCAAGACTGAGGCTGAGAAAGAGGCTCCGTTCA 889  
DB 4007 CTGTCTCAACAATGACCAAAAGCTCTGCAAGACTGAGGCTGAGAAAGAGGCTCCGTTCA 4066  
QY 890 CCTTCATACGACAGTCTGCTGAGAGAGTCTGCTCTCTGAGTCTTCTCTGATCTGACCG 949  
DB 4067 CCTTCATACGACAGTCTGCTGAGAGAGTCTGCTCTCTGAGTCTTCTCTGATCTGACCG 4126  
QY 950 TCAGAGAGCTGGGCAACAAGAGCTCAAGTCAAGAGTGTGTCTTCCCGTTACCTGTTAG 1009  
DB 4127 TCAGAGAGCTGGGCAACAAGAGCTCAAGTCAAGAGTGTGTCTTCCCGTTACCTGTTAG 4186  
QY 1010 TTAGAGGAATCTCCGGGGAACAAGATCCACTTGTCTCTGAGGCGGGAATGTTGAGC 1069  
DB 4187 TTAGAGGAATCTCCGGGGAACAAGATCCACTTGTCTCTGAGGCGGGAATGTTGAGC 4246  
QY 1070 ATCAGAGACACAAGAGTTCGCTGATCATGAAACAACCTGAGGCTGCTCAACAGTCC 1129  
DB 4247 ATCAGAGACACAAGAGTTCGCTGATCATGAAACAACCTGAGGCTGCTCAACAGTCC 4306  
QY 1130 AGGTGCGGCTGGGCTCTCTCATCTGGGTGGCCCTGAGAGTGGACGCTGGTGGGG 1189  
DB 4307 AGGTGCGGCTGGGCTCTCTCATCTGGGTGGCCCTGAGAGTGGACGCTGGTGGGG 4366  
QY 1190 AGAGGTGGCCCTTCAACAAACGCTCAAGGCTGCAAGCGCTTTTGTGTTTATC 1249  
DB 4367 AGAGGTGGCCCTTCAACAAACGCTCAAGGCTGCAAGCGCTTTTGTGTTTATC 4426  
QY 1250 AGCTTCAACCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGGAGGAAGATGTCTGCA 1309

DB 4427 AGCTCACCCCTCGAGGCTGGTCCGCGCTGTCTCAATCTGAGGAAGATGTCTCTGCA 4486  
QY 1310 AGGCTTCTGCGGATGATGCTGTGAGAGGAGTGGAAATGAGAACTGAGTTGATGTTG 1369  
DB 4487 AGGCTTCTGCGGATGATGCTGTGAGAGGAGTGGAAATGAGAACTGAGTTGATGTTG 4546  
QY 1370 ACGACCTCATGATTCAGAGACTCGGGAGTCTGAGCTCCGTGCTCTGTTCACATGACA 1429  
DB 4547 ACGACCTCATGATTCAGAGACTCGGGAGTCTGAGCTCCGTGCTCTGTTCACATGACA 4606  
QY 1430 TCTTCTTCCAGACAGCCACTGTGAGAGTACTACCTTCTTCCACTGATCTCCAGG 1489  
DB 4607 TCTTCTTCCAGACAGCCACTGTGAGAGTACTACCTTCTTCCACTGATCTCCAGG 4666  
QY 1490 ACTTGTGCGGCTTGTACTAGCTGTGAGAGGCGCTGAGAAATGACACGAGCTCTGCGC 1549  
DB 4667 ACTTGTGCGGCTTGTACTAGCTGTGAGAGGCGCTGAGAAATGACACGAGCTCTGCGC 4726  
QY 1550 CTCTGTAGTGAAGAACAAGAGTCTCATGAGCTTAAACAGGCAAGCTTCCATATCC 1609  
DB 4727 CTCTGTAGTGAAGAACAAGAGTCTCATGAGCTTAAACAGGCAAGCTTCCATATCC 4786  
QY 1610 ACTGCTTTGAGTGAAGCGTTCTGTGTTGAGCTCTGTGAGGAAAGAGTGAAGAGCCAC 1669  
DB 4787 ACTGCTTTGAGTGAAGCGTTCTGTGTTGAGCTCTGTGAGGAAAGAGTGAAGAGCCAC 4846  
QY 1670 TGAAGTCTCTGAGGCTGTCGCCGTTCCGCTGGGGGGTGAAGACAGACCTTCCAGTGGG 1729  
DB 4847 TGAAGTCTCTGAGGCTGTCGCCGTTCCGCTGGGGGGTGAAGACAGACCTTCCAGTGGG 4906  
QY 1730 TCTCTGTGTTGGGTCAAGAGCTTAATGCCAACCCAGAGAGACACCTGAGACGCTTCC 1789  
DB 4907 TCTCTGTGTTGGGTCAAGAGCTTAATGCCAACCCAGAGAGACACCTGAGACGCTTCC 4966  
QY 1790 ACTGTCTTTTTCAGACTCAAGACAAGAGTTGTGTCTGCTTGGCATTAACAGCTTCCAAG 1849  
DB 4967 ACTGTCTTTTTCAGACTCAAGACAAGAGTTGTGTCTGCTTGGCATTAACAGCTTCCAAG 5026  
QY 1850 AAGTGTGCTTCGATTAACAGAACTGGAAGCTTGAAGATCTTCTGCTGAGCTCCAGC 1909  
DB 5027 AAGTGTGCTTCGATTAACAGAACTGGAAGCTTGAAGATCTTCTGCTGAGCTCCAGC 5086  
QY 1910 ACTGTCTGATTTTCGAGAAATTCGGGTGATGTCAAGAGGATCTTCCAAAGATAGT 1969  
DB 5087 ACTGTCTGATTTTCGAGAAATTCGGGTGATGTCAAGAGGATCTTCCAAAGATAGT 5146  
QY 1970 CCGCTGAGGCAATGCTGTGCTCTCTGATGATGCGGATTAAGACCTCAATTGAGAGC 2029  
DB 5147 CCGCTGAGGCAATGCTGTGCTCTCTGATGATGCGGATTAAGACCTCAATTGAGAGC 5206  
QY 2030 AGTGGGAAGATTTTGTCTCCATGCTTGGCAACCCACCAACCTGCGGCACTGAGACCTGG 2089  
DB 5207 AGTGGGAAGATTTTGTCTCCATGCTTGGCAACCCACCAACCTGCGGCACTGAGACCTGG 5266  
QY 2090 GCACACGATCTCTGACAGAGGCGGCACTGAAGACCTGTGTGCAACCTGAGGATCCA 2149  
DB 5267 GCACACGATCTCTGACAGAGGCGGCACTGAAGACCTGTGTGCAACCTGAGGATCCA 5326  
QY 2150 CCTGCAAGATACAGACCTGATGTTTGAAGATGCAAGATTAACCTGTGTGTGACAGAC 2209  
DB 5327 CCTGCAAGATACAGACCTGATGTTTGAAGATGCAAGATTAACCTGTGTGTGACAGAC 5386  
QY 2210 TCTGGAAGATGCTATGAGCCACCGTAACTTAAGATCCCTCAACTTGGAGGCAACCCAC 2269  
DB 5387 TCTGGAAGATGCTATGAGCCACCGTAACTTAAGATCCCTCAACTTGGAGGCAACCCAC 5446  
QY 2270 TGAAGAAAGAGATGTAAGATGAGTGAAGGCTTAACCAACCCAAATGTTGTGTTG 2329  
DB 5447 TGAAGAAAGAGATGTAAGATGAGTGAAGGCTTAACCAACCCAAATGTTGTGTTG 5506  
QY 2330 AGTCTTGAAGCTGATGCTGTGATGTTGACCAATGCTGTGACCTGAAGATCTCCAAA 2389

Db 5507 AGCTTTGAGGCTGATGCTGATGATGACCACTGCTGTTACTGAGATCTCCAA 5566  
 Qy 2390 TCCCTTACGACCTCCCGACGCTGAATCTCTGAGCCTTGACAGAAACAAGTGACAGACC 2449  
 Db 5567 TCCCTTACGACCTCCCGACGCTGAATCTCTGAGCCTTGACAGAAACAAGTGACAGACC 5626  
 Qy 2450 AGGAGATATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTTGACAGAACTGA 2509  
 Db 5627 AGGAGATATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTTGACAGAACTGA 5686  
 Qy 2510 TACTGAGAGACTGTGACATCAAGCCAGGTTGCCAGATGCTGACGCTTGCTGCA 2569  
 Db 5687 TACTGAGAGACTGTGACATCAAGCCAGGTTGCCAGATGCTGACGCTTGCTGCA 5746  
 Qy 2570 GCAACCGGAGCTTGACACACTGTGCTTACAAACAAGCTTGAGAAAGGTTAA 2629  
 Db 5747 GCAACCGGAGCTTGACACACTGTGCTTACAAACAAGCTTGAGAAAGGTTAA 5806  
 Qy 2630 ATCTACTGTGATCATGATGAGGCTCCCACTGTATGCTGACAGGCTGATGCTGAATC 2689  
 Db 5807 ATCTACTGTGATCATGATGAGGCTCCCACTGTATGCTGACAGGCTGATGCTGAATC 5866  
 Qy 2690 AGTGCACCTGACACAGGCTGCTGTGTTTCTTGCACTTGCGCTTATGGTTACTCAT 2749  
 Db 5867 AGTGCACCTGACACAGGCTGCTGTGTTTCTTGCACTTGCGCTTATGGTTACTCAT 5926  
 Qy 2750 GGCTGACGACCTGAGCTTACGATGACCTCTGTGAAACAATGCGCTGAAGCTTCTGT 2809  
 Db 5927 GGCTGACGACCTGAGCTTACGATGACCTCTGTGAAACAATGCGCTGAAGCTTCTGT 5986  
 Qy 2810 GCGAGGATGAGAGAACCATCTTGTCACTCCAGGACCTGAGGTTGTAAGTGTCAATC 2869  
 Db 5987 GCGAGGATGAGAGAACCATCTTGTCACTCCAGGACCTGAGGTTGTAAGTGTCAATC 6046  
 Qy 2870 TCACCGCCGCTGCTGTGAGAGTCTGCTGTGTGATCTCGAGAGACACACTGAGAA 2929  
 Db 6047 TCACCGCCGCTGCTGTGAGAGTCTGCTGTGTGATCTCGAGAGACACACTGAGAA 6106  
 Qy 2930 GCTGTGATCTACGACCAATGCCCTGTGAGAGGCTGCGGCTGTGCGAGAGGAC 2989  
 Db 6107 GCTGTGATCTACGACCAATGCCCTGTGAGAGGCTGCGGCTGTGCGAGAGGAC 6166  
 Qy 2990 TGAAGCAAAAGAAAGTGTCTGAGAGAGCTGGGTTGAAGGCAATGTGACCTGCTCG 3049  
 Db 6167 TGAAGCAAAAGAAAGTGTCTGAGAGAGCTGGGTTGAAGGCAATGTGACCTGCTCG 6226  
 Qy 3050 ATTGCTGTGAGGCACTCTCTTGCCCTTCTGCAACCGGATCTGACCACTTAAAC 3109  
 Db 6227 ATTGCTGTGAGGCACTCTCTTGCCCTTCTGCAACCGGATCTGACCACTTAAAC 6286  
 Qy 3110 TGGTGCAAAATTAATCTTCACTGCCAAAGAAATGATGAAAGCTGTGCTGCTTGGCTGTG 3169  
 Db 6287 TGGTGCAAAATTAATCTTCACTGCCAAAGAAATGATGAAAGCTGTGCTGCTTGGCTGTG 6346  
 Qy 3170 CCAAGCTTAACCTTAAGATTAATGG 3194  
 Db 6347 CCAAGCTTAACCTTAAGATTAATGG 6371

RESULT 7  
 AX459891 6939 bp DNA linear PAT 30-NOV-2002  
 LOCUS AX459891 Sequence 34 from Patent WO0240668.  
 DEFINITION AX459891  
 ACCESSION AX459891  
 VERSION AX459891.1 GI:21725654  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS 1 Tschopp, J. and Martinon, F.  
 TITLE Proteins and dna sequences underlying these proteins used for

JOURNAL  
 Parent: WO 0240668-A 34 23-MAY-2002;  
 Apotech Research and Development Ltd. (CH)  
 FEATURES  
 source  
 1..6939  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="NALP13/Fy17.cdna"  
 ORIGIN  
 Query Match 69.5%; Score 2803; DB 6; Length 6939;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 290 AAATTTCACAGCTATGGAACAAGAGGTGCTCACAGCAGCAGACAGACAGAAACAAGAA 349  
 Db 3467 AAATTTCACAGCTATGGAACAAGAGGTGCTCACAGCAGCAGACAGACAGAAACAAGAA 3526  
 Qy 350 TTTTCACAGCTATGGAACAAGAGGTGCTCACAGCAGCAGACAGACAGAAACAAGAACTATG 409  
 Db 3527 TTTTCACAGCTATGGAACAAGAGGTGCTCACAGCAGCAGACAGACAGAAACAAGAACTATG 3586  
 Qy 410 GAGTGACACATGGACTACACAGAGTCAAGTATGACCAAAATTGCTGAGAGAGAGATG 469  
 Db 3587 GAGTGACACATGGACTACACAGAGTCAAGTATGACCAAAATTGCTGAGAGAGAGATG 3646  
 Qy 470 TACGTGATGTTTGAACAACTGCTGTGACATGGCCGAAATGCAACGTTGCTGTG 529  
 Db 3647 TACGTGATGTTTGAACAACTGCTGTGACATGGCCGAAATGCAACGTTGCTGTG 3706  
 Qy 530 CTTTGTGATTCAGACCGGTGGGCTTCCGGCTCCGACGCTGTTTGCACGGAAGTCAAG 589  
 Db 3707 CTTTGTGATTCAGACCGGTGGGCTTCCGGCTCCGACGCTGTTTGCACGGAAGTCAAG 3766  
 Qy 590 GAATTGGGAAATCGGCTCTAGCAAGAGATGCTGTGCTGGGCGAAGGTGACTCT 649  
 Db 3767 GAATTGGGAAATCGGCTCTAGCAAGAGATGCTGTGCTGGGCGAAGGTGACTCT 3826  
 Qy 650 ACCAGGAAATGTTCTCTACGCTTCTTCTCCCGTTAGAGATGACAGGAGAAAG 709  
 Db 3827 ACCAGGAAATGTTCTCTACGCTTCTTCTCCCGTTAGAGATGACAGGAGAAAG 3886  
 Qy 710 AGAGCAGTGTACACAGATTCATCTCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGG 769  
 Db 3887 AGAGCAGTGTACACAGATTCATCTCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGG 3946  
 Qy 770 AGATGATGCCGACCAAGAAAGGCTGTTCATCATTTGACGTTTTGATGACCTGGGCT 829  
 Db 3947 AGATGATGCCGACCAAGAAAGGCTGTTCATCATTTGACGTTTTGATGACCTGGGCT 4006  
 Qy 830 CTGTCTCAACATGACCAAGGCTGTGCAAGACTGAGGCTGAGAGAGAGGCTCCGTTCA 889  
 Db 4007 CTGTCTCAACATGACCAAGGCTGTGCAAGACTGAGGCTGAGAGAGAGGCTCCGTTCA 4066  
 Qy 890 CCTTCATACGAGTGTGCTGAGAAAGTCTGCTGCTGCTGAGTCTTCTGTATGCTGACCG 949  
 Db 4067 CCTTCATACGAGTGTGCTGAGAAAGTCTGCTGCTGCTGAGTCTTCTGTATGCTGACCG 4126  
 Qy 950 TCAGAGAGTGGGACACAGAAAGCTCAAGTCAAGAGTGTGCTTCCCGCTTAACCTTTAG 1009  
 Db 4127 TCAGAGAGTGGGACACAGAAAGCTCAAGTCAAGAGTGTGCTTCCCGCTTAACCTTTAG 4186  
 Qy 1010 TTAGAGAAATCTCCGGGAAACAAGAAATCACTTGTCTTGAAGCGGAGTTGGTGAGC 1069  
 Db 4187 TTAGAGAAATCTCCGGGAAACAAGAAATCACTTGTCTTGAAGCGGAGTTGGTGAGC 4246  
 Qy 1070 ATCAGAGACACAAAGGTTGCTGTGATCATGAAACAACCTGTGAGCTGTGACCAAGTCC 1129  
 Db 4247 ATCAGAGACACAAAGGTTGCTGTGATCATGAAACAACCTGTGAGCTGTGACCAAGTCC 4306  
 Qy 1130 AGTGGCCGCGTGGGCTCTTCACTGCGGCTGCGCTGAGACTGAGAGACCTGTGGGGG 1169



Db 4307 AGGTCCCGCGGGGCTCTCTCATCTGCTGSCCTTGACGCTGAGGACGCTGAGGGG 4366  
Oy 1190 AGAGGTGCCCCCTTCAACGAAGCTCAAGAGCGCTGCAGCGCTTTTGTTGATC 1249  
Db 4367 AGAGGTGCCCCCTTCAACGAAGCTCAAGAGCGCTGCAGCGCTTTTGTTGATC 4426  
Oy 1250 AGCTCACCCCTTGAGAGCGTGTCCGCGCTGTCTCAATCTGAGGAAAGATGTCTGA 1309  
Db 4427 AGCTCACCCCTTGAGAGCGTGTCCGCGCTGTCTCAATCTGAGGAAAGATGTCTGA 4486  
Oy 1310 AGCGTTTGCCCTTATGCTGTGAGAGGAGTGTGAATAGGAAGTCAAGTGTGATG 1369  
Db 4487 AGCGTTTGCCCTTATGCTGTGAGAGGAGTGTGAATAGGAAGTCAAGTGTGATG 4546  
Oy 1370 AGGACCTATGAGTCAAGGACCTCGGAGAGTGTGAGCTCGGTCTGTTTCAATGAACA 1429  
Db 4547 AGGACCTATGAGTCAAGGACCTCGGAGAGTGTGAGCTCGGTCTGTTTCAATGAACA 4606  
Oy 1430 TCCTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCAGTCTCAG 1489  
Db 4607 TCCTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCAGTCTCAG 4666  
Oy 1490 ACTTCTGTGCGGCTTGTACTACTAGTGTGAAGGCGCTGGAATGAGCCAGCTCTGCC 1549  
Db 4667 ACTTCTGTGCGGCTTGTACTACTAGTGTGAAGGCGCTGGAATGAGCCAGCTCTGCC 4726  
Oy 1550 CTCTGTAGTGTGAGAGCAAGAGGTTCATGAGGCTTAAACAGCAGGCTTCAATATCC 1609  
Db 4727 CTCTGTAGTGTGAGAGCAAGAGGTTCATGAGGCTTAAACAGCAGGCTTCAATATCC 4786  
Oy 1610 ACTGCTTTGAGTGAAGGCTTCTTGTGTGAGCTCTGTGAGGAGCAAGCTTGAAGGCCAC 1669  
Db 4787 ACTGCTTTGAGTGAAGGCTTCTTGTGTGAGCTCTGTGAGGAGCAAGCTTGAAGGCCAC 4846  
Oy 1670 TGAAGTCTGCTGTGGGCTGTCCGTTCCCTGTGGGGGTGAAGAGCTTCTGCACTGG 1729  
Db 4847 TGAAGTCTGCTGTGGGCTGTCCGTTCCCTGTGGGGGTGAAGAGCTTCTGCACTGG 4906  
Oy 1730 TCTCTGTGTGGGTGAGAGGCTTAATGACCAACCCAGAGAGCAACCTGTGACGCTTCC 1789  
Db 4907 TCTCTGTGTGGGTGAGAGGCTTAATGACCAACCCAGAGAGCAACCTGTGACGCTTCC 4966  
Oy 1790 ACTGCTTTGAGACTCAAGACAAAGATTTGTGTGCTGTGAGCAATTAACAGCTTCCAG 1849  
Db 4967 ACTGCTTTGAGACTCAAGACAAAGATTTGTGTGCTGTGAGCAATTAACAGCTTCCAG 5026  
Oy 1850 AAGTGTGCTTCCGATTAAACGAACCTGAGCTTGATAGCATCTTCTTCTGCTCCAGC 1909  
Db 5027 AAGTGTGCTTCCGATTAAACGAACCTGAGCTTGATAGCATCTTCTTCTGCTCCAGC 5086  
Oy 1910 ACTGTCGATTTTGGGAAAAATTGGGGTGAATGTCAAGAGGATCTTCCAGAGATGAGT 1969  
Db 5087 ACTGTCGATTTTGGGAAAAATTGGGGTGAATGTCAAGAGGATCTTCCAGAGATGAGT 5146  
Oy 1970 CCGGTGAGGCAATGCTGTGTGCTCTCTATGAGATGCGGGATTAAGACCTCTATGAGAGC 2029  
Db 5147 CCGGTGAGGCAATGCTGTGTGCTCTCTATGAGATGCGGGATTAAGACCTCTATGAGAGC 5206  
Oy 2030 AGTGGAGATTTTGTCTTCAATGCTTGGCACCCACCCACACTGTGCGGAGCTGAGCTGG 2089  
Db 5207 AGTGGAGATTTTGTCTTCAATGCTTGGCACCCACCCACACTGTGCGGAGCTGAGCTGG 5266  
Oy 2090 GCAGAGATCTTGAAGAGGCGGCAATGAAGCCGTGTGTGCCAAGCTGAGGATGCCA 2149  
Db 5267 GCAGAGATCTTGAAGAGGCGGCAATGAAGCCGTGTGTGCCAAGCTGAGGATGCCA 5326  
Oy 2150 CCTGCAAGATACAGACCTGTATGTTTGAAGATGACAGATTAACCTGTGTGAGCACCC 2209  
Db 5327 CCTGCAAGATACAGACCTGTATGTTTGAAGATGACAGATTAACCTGTGTGTGAGCACCC 5386  
Oy 2210 TCTGAGATGTGTGAGGCAACCTGTAACCTTAAGATCCCTCACTTGGAGGCAACCCACC 2269  
Db 5387 TCTGAGATGTGTGAGGCAACCTGTAACCTTAAGATCCCTCACTTGGAGGCAACCCACC 5446

Oy 2270 TGAAGAAAGAGATGTAAGGATGCGGTGAGAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 5447 TGAAGAAAGAGATGTAAGGATGCGGTGAGAGCTTAAACACCCAAATGTTTGTGG 5506  
Oy 2330 AGTCTTGAAGCTGGAATTTGCTGTGATTTGACCCATGCTGTATCTGAAGATCTCCAAA 2389  
Db 5507 AGTCTTGAAGCTGGAATTTGCTGTGATTTGACCCATGCTGTATCTGAAGATCTCCAAA 5566  
Oy 2390 TCTTGAAGACCTCCCGAGCTGAAATCTGTAGGCTGTGGAGGAAACAAGTGAACAACC 2449  
Db 5567 TCTTGAAGACCTCCCGAGCTGAAATCTGTAGGCTGTGGAGGAAACAAGTGAACAACC 5626  
Oy 2450 AGGAGTAAATGCTCTGAGTATGCTTTGAGAGTCTCCAGTGCGCCCTGACAAAGCTGA 2509  
Db 5627 AGGAGTAAATGCTCTGAGTATGCTTTGAGAGTCTCCAGTGCGCCCTGACAAAGCTGA 5686  
Oy 2510 TACTGAGAGCTGTGAGATCAACGCCAGGAGTGTGCAGAGTCTTGAAGCTTCAAGCTCTCA 2569  
Db 5687 TACTGAGAGCTGTGAGATCAACGCCAGGAGTGTGCAGAGTCTTGAAGCTTCAAGCTCTCA 5746  
Oy 2570 GCACCCGAGCTTGAACAACCTGTGCTATTCACAAACACCTGTGGGAAAGAAAGTGTAA 2629  
Db 5747 GCACCCGAGCTTGAACAACCTGTGCTATTCACAAACACCTGTGGGAAAGAAAGTGTAA 5806  
Oy 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTATGTGACAGGCTGATGCTGAATC 2689  
Db 5807 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTATGTGACAGGCTGATGCTGAATC 5866  
Oy 2690 AGTGCACCTGAGACACGCGCTGTGTGTTTCTTGAACCTTGCAGCTTATGAGTAACTCAT 2749  
Db 5867 AGTGCACCTGAGACACGCGCTGTGTGTTTCTTGAACCTTGCAGCTTATGAGTAACTCAT 5926  
Oy 2750 GGCTGAAGCACTGAGGCTTATAGATGAACCTGTGGAAGCAATGAGGTGAAGCTTCTGT 2809  
Db 5927 GGCTGAAGCACTGAGGCTTATAGATGAACCTGTGGAAGCAATGAGGTGAAGCTTCTGT 5986  
Oy 2810 GCGAGTCTATGAGAGAACATCTTGTATCTCCAGACCTGAGATTGATGAAGTGTATC 2869  
Db 5987 GCGAGTCTATGAGAGAACATCTTGTATCTCCAGACCTGAGATTGATGAAGTGTATC 6046  
Oy 2870 TCACCGCGGTGTGTGAGAGTGTCTGTGTGATCTTGAGAGAGACACCTGAAGA 2929  
Db 6047 TCACCGCGGTGTGTGAGAGTGTCTGTGTGATCTTGAGAGAGACACCTGAAGA 6106  
Oy 2930 GCTGTGAATCAAGCAATGCTGTGGTGAAGCTGTGGGTTGCTGCGCTGTGAGAGGAC 2989  
Db 6107 GCTGTGAATCAAGCAATGCTGTGGTGAAGCTGTGGGTTGCTGCGCTGTGAGAGGAC 6166  
Oy 2990 TGAAGCAAAAGACAGTGTCTGACGAGACTCGGTTGAAGGATGTGACTGCTTCTG 3049  
Db 6167 TGAAGCAAAAGACAGTGTCTGACGAGACTCGGTTGAAGGATGTGACTGCTTCTG 6226  
Oy 3050 ATTGTGTGAGGCACTCTCTTGGCCCTTCTGTCAACCGGCATCTGACAGTCTAAACC 3109  
Db 6227 ATTGTGTGAGGCACTCTCTTGGCCCTTCTGTCAACCGGCATCTGACAGTCTAAACC 6286  
Oy 3110 TGTGTGAGATTAATTTAGTGTCCAAAGGATGATGAAGCTGTGTGCGCTTGTGCTGTG 3169  
Db 6287 TGTGTGAGATTAATTTAGTGTCCAAAGGATGATGAAGCTGTGTGCGCTTGTGCTGTG 6346  
Oy 3170 CCAAGCTAACTTACAGATTAATTGG 3194  
Db 6347 CCAAGCTAACTTACAGATTAATTGG 6371

RESULT 8  
AY054986 3885 bp mRNA linear PRI 02-Apr-2002  
LOCUS Homo sapiens maternal-antigen-thal-embryos-require protein (MATER)  
DEFINITION  
ACCESSION AY054986  
VERSION AY054986.1 GI:19882272

KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 3885)
AUTHORS	Tong,Z.-B., Bondy,C.A., Zhou,J. and Nelson,L.M.
TITLE	A human homologue of mouse Mater, a maternal effect gene essential for early embryonic development
JOURNAL	Hum. Reprod. 17 (4), 903-911 (2002)
MEDLINE	21922687
PUBMED	11925379
REFERENCE	2 (bases 1 to 3885)
AUTHORS	Tong,Z.-B., Bondy,C.A. and Nelson,L.M.
TITLE	Direct Submission
JOURNAL	Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD, NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA
FEATURES	Location/Qualifiers
SOURCE	1..3885

gene	1. .3885
.	/gene="WATER"
CDS	1. .3603

/note="oocyte-specific protein; similar to Mus musculus Water encoded by GenBank Accession Numbers AF074018 and AF143559-AF143573"

```
/product="maternal-antigen-that-embryos-require protein
/protein_id="AAL15549.1"
/db_xref="GI:19882273"
```

SSQPCINBEGDKSLTPSSYGLWCLYELDKKEPQTKELLKKKXSESTCSIPQF  
NANVECLALLIHEYGASLAWATSISIFEMNRLRTSEKARDMKRHSPEDEATY  
QGPSKEKYPGISQAVQDSATAETKEQETSQAMEBGATAETBEQETSQAMEQ

FRPRTVLHGSGIGKSALARIIVLCAAGGLYQGMFSYVFPLPVREMRKKESS  
PISREWPSQAPVTEIMSRPELLFIIDGFDLGSVLANDTYLCKDMAEQPPFTT  
SLRKVLLPESFLIVTVRVDGTEKLSKSEVSPRYLLVRGISGEORIHLLERGI

Q L T P R G V A R C L N L E E R V L K R F C R M A E G V M N R K S V F D G D D L M V Q G L G E S E L R A L  
N M I L L P D S H C E E Y T T F H L S L O D F C A L Y Y V L E G L E I B P A L C P L Y E K T R S M E L L E  
G F H I S L M K R F L F G L V S E D V A R P L E V L G C P V P L G V K Q K L M T V S L L G Q P N A T

KGI FPRDESABACPVVPLMMRXTLIEEDMEDFCSMLGTHPHLRQDLGSSILTEE  
KTLCAKLAHPTCKIQTLMFRAOITPGVHLWRIWANRNLBSLNGGTHLKEED  
ACEBALKHPRCLLESRLDCCGLTHACYLKSQILTTPSKLSLAGNKVTDQGV

RSMRLPHCSLQRLMLNQCHLDPAAGCSLALALMGNSWLTHTLSISMNPVEDNGVKLT  
VMREPSCHQDLELVKCHLTAAACEBSLCTISRSRLIKSLDITDNLGDBGVAALG  
LKQNSVLTRLGLKACGLTSDCEBALSTALSCNRHLTSLNVQNNFSPKGMMLCC

IN

Very Match 67.24; Score 2713; DB 9; Length 3885;

290 AAATTTCACAGCTATGGACAAGAGGTCACACAGCAGCAGACAGACAGACAGACAGAAA 340  
 tches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps

566 A A A T T T C A C A A G C T A T G G A A C A A G A A G G T G C C A C A G C A G C A G A C A G A A G A C A A G A A 6  
350 T T T C A C A A G C T A T G G A A C A A G A A G T G C C A C A G C A G C A G A G A C A G A A C A A G C A T G 4

626 TTTCAACGCTATGGAAACAAGAGGTGCCACAGCAGCAGAGACAGAAGAAACAAGACATG 6  
410 GAGGTACACATGGCACTACAAGAGTCACGTGATGACCAATTGCTGAGAGAGAGATG 40

686 GAGGTGACATGGGACTACAGAGTCACTGATGACCAATTCCCTGAGGAGGAGATG 74

QY	470	TACGTCGATGTTGAAAAACATGCTGTGCTGACGTGGCCGGAATGCAACGTTGGCTGTG	529
Db	746	TACGTCGATGTTGAAAAACATGCTGTGCTGACGTGGCCGGAATGCAACGTTGGCTGTG	805
QY	530	CTTTTGATTCGACCCGGTGGGGCTTCCGGGCTCGCACGGTGTCTTCGACCGGAAGTCAG	589
Db	806	CTTTTGATTCAGACCCGGTGGGGCTTCCGGGCTCGCACGGTGTCTTCGACCGGAAGTCAG	865
QY	590	GAATTGGGAAATCGGCTCTAGCCAGGAAGATCGTGTGTGTGGGCGCAGGTGAACTCT	649
Db	866	GAATTGGGAAATCGGCTCTAGCCAGGAAGATCGTGTGTGTGGGCGCAGGTGAACTCT	925
QY	650	ACCAAGGGAATGTCCTCTACGTCTTCTCCCTCCCGTTAGAGAAATCAGCGGAAGAG	709
Db	926	ACCAAGGGAATGTCCTCTACGTCTTCTCCCTCCCGTTAGAGAAATCAGCGGAAGAG	985
QY	710	AGACAGATGTCAACAGATTCAATCCAGGAGTGGCAGACTCCCAAGCTCCGTGACGG	769
Db	986	AGACAGATGTCAACAGATTCAATCCAGGAGTGGCAGACTCCCAAGCTCCGTGACGG	1045
QY	770	AGATCATGTCCGACCAAGAAAGCTGTGTTCATCATTTGACGGTTCGATGACCTGGAGCT	829
Db	1046	AGATCATGTCCGACCAAGAAAGCTGTGTTCATCATTTGACGGTTCGATGACCTGGAGCT	11050
QY	830	CTGTCTCTCAACAAATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAAAGCAGCTCCGTTC	889
Db	1106	CTGTCTCTCAACAAATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAAAGCAGCTCCGTTC	11655
QY	890	CCCTCATACGACGTCTGTGAGGAAGGTCCTGCGCTCCCTGAGTCCCTTCGTATCGTCAACG	949
Db	1166	CCCTCATACGACGTCTGTGAGGAAGGTCCTGCGCTCCCTGAGTCCCTTCGTATCGTCAACG	12255
QY	950	TCAGAGACGTGGGCACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACCTGTTCAG	1009
Db	1226	TCAGAGACGTGGGCACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACCTGTTCAG	12855
QY	1010	TTAAGAGAAATCTCCGGGAAACAAAGATCCAATTGCTCTTGAAGCGGGATTGTTGATGC	1069
Db	1286	TTAAGAGAAATCTCCGGGAAACAAAGATCCAATTGCTCTTGAAGCGGGATTGTTGATGC	1345
QY	1070	ATCGAAGAGACAAAGGGTGGTGGTCGATTCATGTGAACAACGGTAGCTGTGACCAAGTGGC	1129
Db	1346	ATCGAAGAGACAAAGGGTGGTGGTCGATTCATGTGAACAACGGTAGCTGTGACCAAGTGGC	1405
QY	1130	AGGTGCCCGCGTGGGCTCTCTCATCTAGCGTGGCCCTGCAAGCTGACAGACGTGTGGGGG	1189
Db	1406	AGGTGCCCGCGTGGGCTCTCTCATCTAGCGTGGCCCTGCAAGCTGACAGACGTGTGGGGG	1465
QY	1190	AGAGCGTGCCTCTCAACCAACGCTCAACAGGCTGCAACGCCCTTGTGTGTTCATC	1249
Db	1466	AGAGCGTGCCTCTCTCAACCAACGCTCAACAGGCTGCAACGCCCTTGTGTGTTCATC	1525
QY	1250	AGCTCACCCCTCGAGGGGTGTCCGGGCGTGTCTCAATCTGGAAGGAAGATTGTCTCTGA	1309
Db	1526	AGCTCACCCCTCGAGGGGTGTCCGGGCGTGTCTCAATCTGGAAGGAAGATTGTCTCTGA	1585
QY	1310	AGCGCTCTGCGCGTATGAGCTGTGTGAGGAGTGTGAATAGGAATGCAGTGTTGATGATGT	1369
Db	1586	AGCGCTCTGCGCGTATGAGCTGTGTGAGGAGTGTGTGAATAGGAATGCAGTGTTGATGATGT	1645
QY	1370	ACGACCTCATGTGTTCAAGAGACTGGGAGAGTCTGAGCTCCGTGTCTGTGTTTCAATGACA	1429
Db	1646	ACGACCTCATGTGTTCAAGAGACTGGGAGAGTCTGAGCTCCGTGTCTGTGTTTCAATGACA	1705
QY	1430	TCTCTTCTCCAGACAGCCACTGTGAGAGATCTACACTTCTTCCACTCTCAGTCTCCAGG	1489
Db	1706	TCTCTTCTCCAGACAGCCACTGTGAGAGATCTACACTTCTTCCACTCTCAGTCTCCAGG	1765
QY	1490	ACTTCTGTGCGGCTTGTATCTACGTGTTAAGAGGCGCTGGAATGTGACCGACAGCTCTGCGC	1549
Db	1766	ACTTCTGTGCGGCTTGTATCTACGTGTTAAGAGGCGCTGGAATGTGACCGACAGCTCTGCGC	1825
QY	1550	CTCTGTACGTTGAGAGACAAAGGTGCATGTGAGCTTAAACAGGACGCTTCATATCC	1609



Db 1826 CTCTGTACGTTGAGAAAGCAAAAGAGCTCATGAGCTTAAACAGGCAAGCTTCCATATCC 1885  
Qy 1610 ACTGCTTTGATGATGAAGCTTTCTTTGTTGCTGTGAGCGAAGACCTTAAGAGGCCAC 1669  
Db 1886 ACTGCTTTGATGATGAAGCTTTCTTTGTTGCTGTGAGCGAAGACCTTAAGAGGCCAC 1945  
Qy 1670 TGAAGGCTCTGCTGAGGCTGTCCCGTTCCCTGAGGGGTTAAGCAGAAAGCTTCTGCACTGG 1729  
Db 1946 TGAAGGCTCTGCTGAGGCTGTCCCGTTCCCTGAGGGGTTAAGCAGAAAGCTTCTGCACTGG 2005  
Qy 1730 TCTCTGTTGAGGTGAGAGCTTAATGCAACCAACCCAGAGACACCTCTGACGCTTCC 1789  
Db 2006 TCTCTGTTGAGGTGAGAGCTTAATGCAACCAACCCAGAGACACCTCTGACGCTTCC 2065  
Qy 1790 ACTGCTTTTGGAGACTCAAGACAAAGATTGTTGGCTTGGCACTTAACAGCTTCCAG 1849  
Db 2066 ACTGCTTTTGGAGACTCAAGACAAAGATTGTTGGCACTTAACAGCTTCCAG 2125  
Qy 1850 AAGTGTGCTTCCGATTAACAGAACCTGSACTGTATGACATCTTCTTGTGCTCCAGC 1909  
Db 2126 AAGTGTGCTTCCGATTAACAGAACCTGSACTGTATGACATCTTCTTGTGCTCCAGC 2185  
Qy 1910 ACTGTCCGTATTGCGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGATGAGT 1969  
Db 2186 ACTGTCCGTATTGCGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGATGAGT 2245  
Qy 1970 CCGCTGAGGCAATGCTCTGTGCTCTCTATGATGACGGGATTAAGACCTCTATGAGAGC 2029  
Db 2246 CCGCTGAGGCAATGCTCTGTGCTCTCTATGATGACGGGATTAAGACCTCTATGAGAGC 2305  
Qy 2030 AAGTGGAGAGATTTGCTCTCATGCTTGGCAACCAACCACTGTGAGGAGCTGAGCTGG 2089  
Db 2306 AAGTGGAGAGATTTGCTCTCATGCTTGGCAACCAACCACTGTGAGGAGCTGAGCTGG 2365  
Qy 2090 GCAGCAGCATCTCTGACAGAGCGGCGCATGAGAACCTCTGTGTGCAAGCTGAGGATCCCA 2149  
Db 2346 GCAGCAGCATCTCTGACAGAGCGGCGCATGAGAACCTCTGTGTGCAAGCTGAGGATCCCA 2425  
Qy 2150 CCTGCAATATACAGACCTCTGATGTTAAGAAATGACCAATTAACCTCTGTGTGCAAGC 2209  
Db 2426 CCTGCAATATACAGACCTCTGATGTTAAGAAATGACCAATTAACCTCTGTGTGCAAGC 2485  
Qy 2210 TCTGAGAAATGCTCATGAGCAACCGTAACTTAAGTCCCTCAACTTGGAGGACACCCAC 2269  
Db 2486 TCTGAGAAATGCTCATGAGCAACCGTAACTTAAGTCCCTCAACTTGGAGGACACCCAC 2545  
Qy 2270 TGAAGGAGAGATGTAAGATGAGCGGTGTAAGCTTAAACAACCAAAATGTTTGTGG 2329  
Db 2546 TGAAGGAGAGATGTAAGATGAGCGGTGTAAGCTTAAACAACCAAAATGTTTGTGG 2605  
Qy 2330 AGTCTTTGAGGCTGATGCTGTGATGTAAGCAACCTGCTTAACTGTAAGATCTTCCCAA 2389  
Db 2606 AGTCTTTGAGGCTGATGCTGTGATGTAAGCAACCTGCTTAACTGTAAGATCTTCCCAA 2665  
Qy 2390 TCTCTACGACCTCCCCAGACCTGAATCTCTGAGGCTGTGAGAGAAACAAGGTGACAGC 2449  
Db 2666 TCTCTACGACCTCCCCAGACCTGAATCTCTGAGGCTGTGAGAGAAACAAGGTGACAGC 2725  
Qy 2450 AAGGAGTAATGCTCTGAGTATGCTTGAAGTCTCCAGTGCCTCTGCAAGAGCTGA 2509  
Db 2726 AAGGAGTAATGCTCTGAGTATGCTTGAAGTCTCCAGTGCCTCTGCAAGAGCTGA 2785  
Qy 2510 TACTGAGAGACTGTGAGCAACAGCAGCGGTGCAAGATGTCGAGCTTCCAGCCTCGCA 2569  
Db 2786 TACTGAGAGACTGTGAGCAACAGCAGCGGTGCAAGATGTCGAGCTTCCAGCCTCGCA 2845  
Qy 2570 GCAACCGGAGCTTGACACACCTGCTATCCAAACAAGCCTGTGGAGAACGAAGGTGTA 2629  
Db 2846 GCAACCGGAGCTTGACACACCTGCTATCCAAACAAGCCTGTGGAGAACGAAGGTGTA 2905  
Qy 2630 ATCTACTGTGTGATCATGAGGCTTCCCACTGATGCTGTGAGAGGCTGATGCTGAATC 2689

Db 2906 ATCTACTGTGTGATCATGAGGCTTCCCACTGATGCTGTGAGAGGCTGATGCTGAATC 2965  
Qy 2690 AAGTGCACCTGGAGACAGGCTGTGAGTGTCTTCTGACCTTATGAGGTAATCAT 2749  
Db 2966 AAGTGCACCTGGAGACAGGCTGTGAGTGTCTTCTGACCTTATGAGGTAATCAT 3025  
Qy 2750 GGTCTGACGCACTTGAAGCTTATGATGAACCTGTGAGAAACAATGAGCGTGAAGCTTCTGT 2809  
Db 3026 GGTCTGACGCACTTGAAGCTTATGATGAACCTGTGAGAAACAATGAGCGTGAAGCTTCTGT 3085  
Qy 2810 GCGAGTATGATGAAGAACATCTTGTATCTCCAGAACCTGTGAGATGTTGTAAGTGTATC 2869  
Db 3086 GCGAGTATGATGAAGAACATCTTGTATCTCCAGAACCTGTGAGATGTTGTAAGTGTATC 3145  
Qy 2870 TCACGCGCGGTGTGATGAGAGTGTCTGATGATCTCGAGAGAGACACCTGAAGA 2929  
Db 3146 TCACGCGCGGTGTGATGAGAGTGTCTGATGATCTCGAGAGAGACACCTGAAGA 3205  
Qy 2930 GCTGTGATCTTCAACGGAATGCTCCCTGGGTGAACGCTGAGGCTGTGCGAGGAGC 2989  
Db 3206 GCTGTGATCTTCAACGGAATGCTCCCTGGGTGAACGCTGAGGCTGTGCGAGGAGC 3265  
Qy 2990 TGAAGCAAAAGACAGTGTCTGACAGAGACTCGGCTTGAAGGCAATGTGACTGACTTCTG 3049  
Db 3266 TGAAGCAAAAGACAGTGTCTGACAGAGACTCGGCTTGAAGGCAATGTGACTGACTTCTG 3325  
Qy 3050 ATTGCTGTGAGGCACTCTCTGAGCCTTTCCTGCAACCGGCACTGTGACAGCTTAAC 3109  
Db 3326 ATTGCTGTGAGGCACTCTCTGAGCCTTTCCTGCAACCGGCACTGTGACAGCTTAAC 3385  
Qy 3110 TGTGTGCAAAATTAATCTTCAAGTCCCAAGAAATGATGAAGCTGTGTTGAGCTTGTCTGTC 3169  
Db 3386 TGTGTGCAAAATTAATCTTCAAGTCCCAAGAAATGATGAAGCTGTGTTGAGCTTGTCTGTC 3445  
Qy 3170 CCAAGTCTTAATTAATTAATGAGGCTGTGAGAAATGAGCACTGCTGTGCAAAATTAAGA 3229  
Db 3446 CCAAGTCTTAATTAATTAATGAGGCTGTGAGAAATGAGCACTGCTGTGCAAAATTAAGA 3505  
Qy 3230 AGCTGTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGAATGAGCGTATGAGGAT 3289  
Db 3506 AGCTGTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGAATGAGCGTATGAGGAT 3565  
Qy 3290 CTTTGTGATGAAGTGAACCG 3308  
Db 3566 CTTTGTGATGAAGTGAACCG 3584

RESULT 9  
AX427610  
LOCUS AX427610 3900 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 23 from Patent WO232955.  
ACCESSION AX427610  
VERSION AX427610.1 GI:21537730  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
Nelson, L.M. and Tong, Z.-B.  
Human gene critical to fertility  
Patent: WO 0232955-A 23 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
FEATURES  
location/Qualifiers  
CDS  
1..3900  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..3603  
/note="unamed protein product"  
/codon\_start=1  
/protein\_id="CAD35973.1"  
/db\_xref="GI:21537731"

/translation="MKVAGLELGAALLASPRALVTLSTPTCSILPKNLPONT  
SSOPCIKMEGDSLFSSYGLQMCLEYLDEKEFORFKLKKSSSSTSCSPOREIE  
NANVECLALLHEYYGASLAMAATSIIPENNMLRTLSKAPRDMKSSREDEAT  
QPSKEXKPGISQAOVQDSATTAETKEDEISQAMEDEGATTAETBEDEISQAMEDEGA  
TAETBEDEGSDWYKSHVMTKFAEEDVRSFENTADPEMOTLAKESDRWG  
FPRRTVVLKSGIKSALARIIVLCAEYVGFVPLPVREMOCKESVTE  
FISREMPDSQAVTEIMSRPERLFIIDFDGLGVNLNDITSCOMKHEDEBEHO  
SLRRVLLPESFLIVTRDVTEKLSKESVSPRYLLNGISEORHILLEGIEHO  
KTQGRALINNEELIDCOVPAVGLICVALODPVGSVAPRNOTITGLHAPAFH  
QUTPGVVRCLNLEBRVLRKPCMAVEGWNKRSVFGDLMVQGLSESLRSLFPH  
NMLLPDSHEEYTFPHLSQDFCALVYVLEGEIBALCPVLEKRSMEKOA  
GHIHSLMKRFLFGLVSEDEVRLPVLGCVPLGVKOKLHWSLLGQCPNLTTPG  
DTLDAPLCLFETODEFVRLALNSFOEVLPINONLDLASSFCOHQCYKINADV  
KGIIPRDSAEACPVPLMMPKXLLIEMEDPCSMGLGPHPLRDLGSSITLBERAM  
KTLCAKLRPKTIOTIMPRNQTITPVCHLIRYMANRNLASLNGVHLEEDVRM  
ACBALNHPKCLSLRLDCCGLTHACVYLISQILTPSPLKSLSLAGNKVTDQVPL  
SDALRVSQDQKLILEDCGLTATGCSLASLAVNSRLTHLCLNNSISGNNVILC  
RSMRLPHCSLQRLMNOCHLDIAGCSLALMNGNSLTHLSLNNPVEDNOCPLICE  
VWRBPSCHLODELEVKCHLTAACCSLSVSRSHKSLIDLVDNADGQVVALCEG  
LKORNSVLRGLKAGLSDCCALSLALSCNRHLTSINNTYNNPSPRGMKLSAF  
ACPSINLOITIGMKQVPOVQIRKLIEVQLPRVYIDGSHSFDDEDRYMKXN"

## ORIGIN

Query Match 67.2%; Score 2713; DB 6; Length 3900;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

290 AATTTTCAACAGCTATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAA 349  
566 AATTTTCAACAGCTATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAA 625  
350 TTTCAAGCTATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAAACAAG 409  
626 TTTCAAGCTATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAAACAAG 685  
410 GAGGTGACACATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAGATG 469  
666 GAGGTGACACATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAGATG 745  
470 TAGCGTGAATTTTGAACAACCTGCTGCTGACGTGCGCGGAAATGCAACGTTGCTG 529  
746 TAGCGTGAATTTTGAACAACCTGCTGCTGACGTGCGCGGAAATGCAACGTTGCTG 805  
530 CTTTGTGATTTGACACCGGTGGGGCTTCGGGCTTCGCAACGCGTGTTCGCAACGGA 589  
806 CTTTGTGATTTGACACCGGTGGGGCTTCGGGCTTCGCAACGCGTGTTCGCAACGGA 865  
590 GAAATTGGGAAATGCGCTTACGCAAGAGATGCTGCTGCTGCGGCGCAAGGTGAGACT 649  
866 GAAATTGGGAAATGCGCTTACGCAAGAGATGCTGCTGCTGCGGCGCAAGGTGAGACT 925  
650 ACGAGGAAATGCTTCTCTAAGCTCTTCTCTCTCCCGTTAGAGAGATGCAAGCGGAAG 709  
926 ACGAGGAAATGCTTCTCTAAGCTCTTCTCTCTCCCGTTAGAGAGATGCAAGCGGAAG 985  
710 AGAGAGATGCTCAAGAGTTCATCTCAGAGAGATGCGCAGACTTCGCAAGCGGTGAGCG 769  
986 AGAGAGATGCTCAAGAGTTCATCTCAGAGAGATGCGCAGACTTCGCAAGCGGTGAGCG 1045  
770 AGATCATGCTCCGACCAAGAAAGCTGTTGTTGATCATTTGAGCGGTTTCAGATGAGCT 829  
1046 AGATCATGCTCCGACCAAGAAAGCTGTTGTTGATCATTTGAGCGGTTTCAGATGAGCT 1105  
830 CTGTCTTCAACATGACCAAAAGCTCTGCAAAAGCTGAGGCTGAGAGAGACGCTCCGTTCA 889  
1106 CTGTCTTCAACATGACCAAAAGCTCTGCAAAAGCTGAGGCTGAGAGAGACGCTCCGTTCA 1165  
890 CCTCATATGCGAGTCTGCTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949  
1166 CCTCATATGCGAGTCTGCTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225  
950 TCAGAGAGTGGGACAGAGAGATCAAGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1009

1226 TCAGAGAGTGGGACAGAGAGATCAAGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1285  
1010 TTAGAGAAATCTCCGGGGGAAACAAGATTCATCTGCTCTTGAAGCGGAGATTGGTGAC 1069  
1286 TTAGAGAAATCTCCGGGGGAAACAAGATTCATCTGCTCTTGAAGCGGAGATTGGTGAC 1345  
1070 ATCAGAGACCAAGAGTGGGTGATCAGAAACAACGAGAGTGTGCTGCTGCTGCTGCTGCTGCT 1129  
1346 ATCAGAGACCAAGAGTGGGTGATCAGAAACAACGAGAGTGTGCTGCTGCTGCTGCTGCTGCT 1405  
1130 AGGTCCCGCGGTGGGCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189  
1406 AGGTCCCGCGGTGGGCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1465  
1190 AGAGCTGGCGCCCTTCAACAACAGCTCAAGCTGCAAGCGCGCTTTGTGTTCAATC 1249  
1466 AGAGCTGGCGCCCTTCAACAACAGCTCAAGCTGCAAGCGCGCTTTGTGTTCAATC 1525  
1250 AGCTCAACCCCTGAGAGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1309  
1526 AGCTCAACCCCTGAGAGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1585  
1310 AGCGCTTGGCGGTATGCTGCTGAGAGAGTGTGGAATAGGAAGTCACTGTTGATGCTG 1369  
1586 AGCGCTTGGCGGTATGCTGCTGAGAGAGTGTGGAATAGGAAGTCACTGTTGATGCTG 1645  
1370 AGACCTATGATGTTCAAGAGACTCGGGAGTGTGAGCTCGTGTCTGTTTCAATGAACA 1429  
1646 AGACCTATGATGTTCAAGAGACTCGGGAGTGTGAGCTCGTGTCTGTTTCAATGAACA 1705  
1430 TCCCTTCCCGAGAGCGACCTGTAAGAGTCTACACTTCTTCCACTGCTGCTGCTGCTGCTGCTG 1489  
1706 TCCCTTCCCGAGAGCGACCTGTAAGAGTCTACACTTCTTCCACTGCTGCTGCTGCTGCTGCTG 1765  
1490 ACTTCTGTCGCGCTTGTACTAGAGTGTGAGAGCGCTGGAATGAGCAGCTCTCTGCC 1549  
1766 ACTTCTGTCGCGCTTGTACTAGAGTGTGAGAGCGCTGGAATGAGCAGCTCTCTGCC 1825  
1550 CTCTGATGTTGAGAGCAAAAGAGTTCATGAGCTTAAACAGGACGCTTCAATATCC 1609  
1826 CTCTGATGTTGAGAGCAAAAGAGTTCATGAGCTTAAACAGGACGCTTCAATATCC 1885  
1610 ACTGCTTGTGATGAGAGCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1669  
1886 ACTGCTTGTGATGAGAGCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1945  
1670 TGAAGTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729  
1946 TGAAGTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2005  
1730 TCTCTGTTGAGGCTCAGAGCTTAATGCAACACCCAGAGACACCTTGAAGCGCTTCC 1789  
2006 TCTCTGTTGAGGCTCAGAGCTTAATGCAACACCCAGAGACACCTTGAAGCGCTTCC 2065  
1790 ACTGCTTGTGAGACTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1849  
2066 ACTGCTTGTGAGACTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2125  
1850 AAGTGTGCTTCCGATTAACAGAACTGAGCTTGAATGATCTTCTGCTGCTGCTGCTGCTGCTGCT 1909  
2126 AAGTGTGCTTCCGATTAACAGAACTGAGCTTGAATGATCTTCTGCTGCTGCTGCTGCTGCTGCT 2185  
1910 ACTGTCGATTTTGGGAAATTCGGGTGATGTCAGAGGATCTTCCCAAGAGATGAGT 1969  
2186 ACTGTCGATTTTGGGAAATTCGGGTGATGTCAGAGGATCTTCCCAAGAGATGAGT 2245  
1970 CCGCTGAGGAGATGTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2029  
2246 CCGCTGAGGAGATGTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2305  
2030 AGTGGAGAAATTTTGTGCTCAATGCTTGGAGCCAGCCACCTGCGGAGCTGAGACCTGG 2089  
2306 AGTGGAGAAATTTTGTGCTCAATGCTTGGAGCCAGCCACCTGCGGAGCTGAGACCTGG 2365

OY	2090	GCAGACAGATCTCTGCACAAAGCGGGCGATGAAAGACCTCTGTGGCCAAAGCTGAGGCATCCCA	2149
Db	2366	GCAGACAGATCCTGCACAAAGCGGGCGCAATGAAGACCTCTGTGGCCAAAGCTGAGGCATCCCA	2425
OY	2150	CCTGCAAAGTACAGACCCCTGATGTTTAAATAATGCACAGATTACCCCTGTGTGCAGCAC	2209
Db	2426	CCTGCAAAGTACAGACCCCTGATGTTTAAATAATGCACAGATTACCCCTGTGTGCAGCAC	2485
OY	2210	TCTGGAGATGCTCATGSCCAACCGTAACCTTAAGTCCCTTCACTTGGGAGCACCCAC	2269
Db	2486	TCTGGAGATGCTCATGSCCAACCGTAACCTTAAGTCCCTTCACTTGGGAGCACCCAC	2545
OY	2270	TGAAGGAAGAGATTTAAGAGATGGGAGTGAAGCCCTTAAACACCCAAATAGTTGTTGG	2329
Db	2546	TGAAGGAAGAGATTTAAGAGATGGGAGTGAAGCCCTTAAACACCCAAATAGTTGTTGG	2605
OY	2330	AGTCTTTGAGGCTGATGCTGTGGATTGACCCATGCTCTGTACTGTGAAGATCTCCCAA	2389
Db	2606	AGTCTTTGAGGCTGATGCTGTGGATTGACCCATGCTCTGTACTGTGAAGATCTCCCAA	2665
OY	2330	TCCTTACGACCTCCCCAGCCTGAATCTGTAGCCTGGCAGAAACAAGTGCAGACC	2449
Db	2666	TCCTTACGACCTCCCCAGCCTGAATCTGTAGCCTGGCAGAAACAAGTGCAGACC	2725
OY	2450	AGGAGTAACTGCTCTCAGGTATGCTTGAAGTCCCAATGGGCCCTGAGAGAAAGCTGA	2509
Db	2726	AGGAGTAACTGCTCTCAGGTATGCTTGAAGTCCCAATGGGCCCTGAGAGAAAGCTGA	2785
OY	2510	TACTGAGAGACTGTGGCATCAACGCAACGGGTTGCCAAGTCTGGCCCTCAGCCCTCGTCA	2569
Db	2786	TACTGAGAGACTGTGGCATCAACGCAACGGGTTGCCAAGTCTGGCCCTCAGCCCTCGTCA	2845
OY	2570	GCAACCGGAGCTTGCACAACCTGTGCTTATCCACACAAGCCTTGGGAGCAGAAGTGTAA	2629
Db	2846	GCAACCGGAGCTTGCACAACCTGTGCTTATCCACACAAGCCTTGGGAGCAGAAGTGTAA	2905
OY	2630	ATCTACTGTGTCGAATCCAAATGAGGCTTCCCACTGAAGTCCAGAGCTGATGTCGAATC	2689
Db	2906	ATCTACTGTGTCGAATCCAAATGAGGCTTCCCACTGAAGTCCAGAGCTGATGTCGAATC	2965
OY	2690	AGTGCACCTGGAACAACCGCTGGCTGTGTTCTTTCGACTTGGCGTTATAGGATTAATCAT	2749
Db	2966	AGTGCACCTGGAACAACCGCTGGCTGTGTTCTTTCGACTTGGCGTTATAGGATTAATCAT	3025
OY	2750	GGCTGACGCACTGAGCCTTAGCATGAACCTCTGTGGAAACAATAGCCTGAAGCTTCTGT	2809
Db	3026	GGCTGACGCACTGAGCCTTAGCATGAACCTCTGTGGAAACAATAGCCTGAAGCTTCTGT	3085
OY	2810	GGAGAGTATGAGGAACAATCTTGTATCTCCAGAACCTGAGTGGTAAAGTATC	2869
Db	3086	GGAGAGTATGAGGAACAATCTTGTATCTCCAGAACCTGAGTGGTAAAGTATC	3145
OY	2870	TCACCGCCGCGTGTGAGAGAGTCTGTCTGTGTATCTCGAGAGCAGACACTGTGAAGA	2929
Db	3146	TCACCGCCGCGTGTGAGAGAGTCTGTCTGTGTATCTCGAGAGCAGACACTGTGAAGA	3205
OY	2930	GCCTGGAATCTCAACGACCAATGCGCTGGGTGAACGGTGGGTTGCTGCGCTTGGCAGAGGAC	2989
Db	3206	GCCTGGAATCTCAACGACCAATGCGCTGGGTGAACGGTGGGTTGCTGCGCTTGGCAGAGGAC	3265
OY	2990	TGAAGCAAAAGAACAGTGTCTGACGAACCTCGGTTGAAGGCAATGTGACATGACCTTCTG	3049
Db	3266	TGAAGCAAAAGAACAGTGTCTGACGAACCTCGGTTGAAGGCAATGTGACATGACCTTCTG	3325
OY	3050	ATTGCTGTGAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCAAGTCTTAAAC	3109
Db	3326	ATTGCTGTGAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCATCTGACCAAGTCTTAAAC	3385
OY	3110	TGTGTACAAATTAATCTTCAGTCCCAAGGAATGATGAAGCTGTGTTGGCCCTTGTGCTGTG	3169
Db	3386	TGTGTACAAATTAATCTTCAGTCCCAAGGAATGATGAAGCTGTGTTGGCCCTTGTGCTGTG	3445

Qy	3170	CCAGCTCACTTACAGATTAATTTGGGCGTGTGGAAATGCACTACCTGTGCAATTAAGA	3222
Db	3446	CCAGCTCACTTACAGATTAATTTGGGCGTGTGGAAATGCACTACCTGTGCAATTAAGA	3505
Qy	3230	AGCTCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGTGTAATTCAGCGTATTTGGCATTT	3289
Db	3506	AGCTCTGAGGAAGTGCAGCTACTCAAGCCCGAGTGTGTAATTCAGCGTATTTGGCATTT	3565
Qy	3290	CTTTGATGAAGATGACCG 3308	
Db	3566	CTTTGATGAAGATGACCG 3584	
RESULT 10			
LOCUS	CQ731113	2753 bp	linear
DEFINITION	Sequence 17047 from Patent WO02068579.		
ACCESSION	CQ731113		
VERSION	CQ731113.1	GI:42306827	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
TITLE	Kites, such as nucleic acid arrays, comprising a majority of humanecons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A, 17047 06-SEP-2002;		
FEATURES			
source	1..2753	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	48.0%; Score 1936; DB 6; Length 2753;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 1936; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	404	GACATGAGGTGACACATGGGACTACACAGAATCACAGTATGACCAAAATTCGCTGAGGAGG	463
Db	1	GACATGAGGTGACACATGGGACTACACAGAATCACAGTATGACCAAAATTCGCTGAGGAGG	60
Qy	464	AGGATGTACGTGCTGATTTGAAAAACATGCTGCTGACCTGGCCGGAAATGCAACGTTGG	523
Db	61	AGGATGTACGTGCTGATTTGAAAAACATGCTGCTGACCTGGCCGGAAATGCAACGTTGG	120
Qy	524	CTGTGTCTTTTGATTCACAGCCGTGCGGCTTCGCGACCGTGTCTGACACGAA	583
Db	121	CTGTGTCTTTTGATTCACAGCCGTGCGGCTTCGCGACCGTGTCTGACACGAA	180
Qy	584	AGTCAGAAATTTGGAAATCGGCTTACGACAGAGATGCTGTGCTGGCGCAAGTG	643
Db	181	AGTCAGAAATTTGGAAATCGGCTTACGACAGAGATGCTGTGCTGGCGCGAAAGTG	240
Qy	644	GACTTACCAAGGAATGTTCTCTCAACGCTTCTTCCCTCCCGTTAAGAGATGACAGGGA	703
Db	241	GACTTACCAAGGAATGTTCTCTCAACGCTTCTTCCCTCCCGTTAAGAGATGACAGGGA	300
Qy	704	AGAAGGAGAGAGTGTCAACAGATTCAATCTCAAGGAGTGGCCAGACTCCAGGCTCCGG	763
Db	301	AGAAGGAGAGAGTGTCAACAGATTCAATCTCAAGGAGTGGCCAGACTCCAGGCTCCGG	360
Qy	764	TGACGGAATCATGTCGCCACAGAAAGGCTGTGTTTCATCATTTGACGATTTGCATGACC	823
Db	361	TGACGGAATCATGTCGCCACAGAAAGGCTGTGTTTCATCATTTGACGATTTGCATGACC	420
Qy	824	TGGGCTCTGTCTCAACATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGAGGCTC	883
Db	421	TGGGCTCTGTCTCAACATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGAGGCTC	480

QY 884 CGTTGACCTCTAATAGCAGTCTGAGAGAGAGTCTGCTCCCTGAGTCTTCTGATCG 943  
 DB 481 CGTTACCTCTAATAGCAGTCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATCG 540  
 QY 944 TCACGTCAGAGAGTGGGCAAGAGAGTCAAGTCAAGAGTCTGCTCCCTGATCC 1003  
 DB 541 TCACGTCAGAGAGTGGGCAAGAGAGTCAAGTCAAGAGTCTGCTCCCTGATCC 600  
 QY 1004 TGTAGTGAAGAAATCTCCGGGAAACAAAGATCACTTGTCTCTTGAAGCGGGATTTG 1063  
 DB 601 TGTAGTGAAGAAATCTCCGGGAAACAAAGATCACTTGTCTCTTGAAGCGGGATTTG 660  
 QY 1064 GTGAGCATCAAGAAACAAAGAGTGGGCAAGAGTCAAGTCAAGAGTCTGCTGAGC 1123  
 DB 661 GTGAGCATCAAGAAACAAAGAGTGGGCAAGAGTCAAGTCAAGAGTCTGCTGAGC 720  
 QY 1124 AGTGCAGAGTGGGCGCGGCTCTCATCTGAGTGGGCGCTGAGTGCAGAGAGTGG 1183  
 DB 721 AGTGCAGAGTGGGCGCGGCTCTCATCTGAGTGGGCGCTGAGTGCAGAGAGTGG 780  
 QY 1184 TGGGGAGAGCGTGGCGCGCTTCAACAAAGCTCAAGAGCTGCAAGCGCGCTTTTGTGT 1243  
 DB 781 TGGGGAGAGCGTGGCGCGCTTCAACAAAGCTCAAGAGCTGCAAGCGCGCTTTTGTGT 840  
 QY 1244 TTTCATCAGTCAACCTCTGAGAGCGTGGTCCGCGCTGTCTCATCTGAGAGAAAGTTG 1303  
 DB 841 TTTCATCAGTCAACCTCTGAGAGCGTGGTCCGCGCTGTCTCATCTGAGAGAAAGTTG 900  
 QY 1304 TCCCTAAGCGCTTCTGCGGTATGAGTGGGAGAGTGGAAATGGAAGTCAAGTTTG 1363  
 DB 901 TCCCTAAGCGCTTCTGCGGTATGAGTGGGAGAGTGGAAATGGAAGTCAAGTTTG 960  
 QY 1364 ATGAGTGAAGACCTCAATGTTTCAAGAGTCTGGGAGTCTGAGTCTGTTTCAACA 1423  
 DB 961 ATGAGTGAAGACCTCAATGTTTCAAGAGTCTGGGAGTCTGAGTCTGTTTCAACA 1020  
 QY 1424 TGAACATCTTCTCCAGACAGCCACTGTGAGAGTCAACCTTCTTCCACTCAAGTC 1483  
 DB 1021 TGAACATCTTCTCCAGACAGCCACTGTGAGAGTCAACCTTCTTCCACTCAAGTC 1080  
 QY 1484 TCCAGAGCTTCTGAGCGCGCTTGTACTAGTGTGAAGGCGTGGAAATGCGCAGAGTC 1543  
 DB 1081 TCCAGAGCTTCTGAGCGCGCTTGTACTAGTGTGAAGGCGTGGAAATGCGCAGAGTC 1140  
 QY 1544 TCTGCCCTCTGTACGTTGAGAGAGAGAGTTCATGAGCTTAAACAGCAGGCTTCC 1603  
 DB 1141 TCTGCCCTCTGTACGTTGAGAGAGAGAGTTCATGAGCTTAAACAGCAGGCTTCC 1200  
 QY 1604 ATATCACTCGCTTGTGATGAAGCGTTTCTTGTGGCCTGTGAGCGGAAGCTTGAAGA 1663  
 DB 1201 ATATCACTCGCTTGTGATGAAGCGTTTCTTGTGGCCTGTGAGCGGAAGCTTGAAGA 1260  
 QY 1664 GGGCACTGAGAGTCTGAGTGGGCTGCTCCGTTCCCTGGGGGTGAAGAGCTTGTGC 1723  
 DB 1261 GGGCACTGAGAGTCTGAGTGGGCTGCTCCGTTCCCTGGGGGTGAAGAGCTTGTGC 1320  
 QY 1724 ACTGGGCTCTCTGTGGGTGAGAGAGCTTATGCAACCAACCCAGAGAGACCTTGAAG 1783  
 DB 1321 ACTGGGCTCTCTGTGGGTGAGAGAGCTTATGCAACCAACCCAGAGAGACCTTGAAG 1380  
 QY 1784 CTTTCACTGTCTTTCGAGACTCAAGACAAAGATTTGCTTGGCATTAAACAGCT 1843  
 DB 1381 CTTTCACTGTCTTTCGAGACTCAAGACAAAGATTTGCTTGGCATTAAACAGCT 1440  
 QY 1844 TCCAGAGAGTGGGCTTCCGATTAACCAAGACCTGAGCTTGTATGATCTTCTTCTGCC 1903  
 DB 1441 TCCAGAGAGTGGGCTTCCGATTAACCAAGACCTGAGCTTGTATGATCTTCTTCTGCC 1500  
 QY 1904 TCCAGACCTGTCCGATTTTTCGAGAAATTCGGGTGAGTGTCAAGAGGATCTTCCCAAG 1963  
 DB 1501 TCCAGACCTGTCCGATTTTTCGAGAAATTCGGGTGAGTGTCAAGAGGATCTTCCCAAG 1560

QY 1964 ATGATCCGCTGAGGAGATGCTGAGGCTGAGGCTGATGAGTGCAGGAGATGAGACCTCATTTG 2023  
 DB 1561 ATGATCCGCTGAGGAGATGCTGAGGCTGAGGCTGATGAGTGCAGGAGATGAGACCTCATTTG 1620  
 QY 2024 AGAGCAGTGGAGAGATTTCTGCTCATGCTTGGCAACCAACCAACCTTGGGAGCTGG 2083  
 DB 1621 AGAGCAGTGGAGAGATTTCTGCTCATGCTTGGCAACCAACCAACCTTGGGAGCTGG 1680  
 QY 2084 ACCTGGGAGAGACATCTTGAAGAGCGGCGCATGAGAGACCTTGTGTCCAAAGTGAAGC 2143  
 DB 1681 ACCTGGGAGAGACATCTTGAAGAGCGGCGCATGAGAGACCTTGTGTCCAAAGTGAAGC 1740  
 QY 2144 ATCCACCTGAGATACAGACCCGATGTTAGAAATGACAGATTAACCTCCGTGTGTC 2203  
 DB 1741 ATCCACCTGAGATACAGACCCGATGTTAGAAATGACAGATTAACCTCCGTGTGTC 1800  
 QY 2204 AGCAGCTTGAAGATCTGATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGAGCA 2263  
 DB 1801 AGCAGCTTGAAGATCTGATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGAGCA 1860  
 QY 2264 CCCACCTGAAGAGAGATGTAAGATGCGGTGTAAGCTTAAACCAACCAATGTT 2323  
 DB 1861 CCCACCTGAAGAGAGATGTAAGATGCGGTGTAAGCTTAAACCAACCAATGTT 1920  
 QY 2324 TGTGGAGCTTTGAG 2339  
 DB 1921 TGTGGAGCTTTGAG 1936

## RESULT 11

AC011470/c

LOCUS AC011470 157141 bp DNA linear PRI 15-JUL-2000  
 DEFINITION Homo sapiens chromosome 19 clone CTC-490M10, complete sequence.

AC011470

VERSION AC011470.5 GI:9211204  
 KEYWORDS HTG.

SOURCE Homo sapiens (human)  
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 157141)  
 DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 157141)  
 DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 157141)  
 DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA

COMMENT On Jul 15, 2000 this sequence version replaced gi:7630109.  
 Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov  
 www.sbgc.stanford.edu

QUALITY: Phrap Quality >=40 99.9% of Sequence;  
 Estimated Total Number of Errors is 0.1.

FEATURES  
 source 1. 157141

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="CTC-490M10"

ORIGIN  
 Query Match 38.4%; Score 1549; DB 9; Length 157141;  
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 1599; Conservative 0;

QY	1482	CTTCAGAGACTCTGTCGCGCCCTTGATACACGCTTGAAGAGGGCTCTGAATGAGCCAC	1541
Db	40921	TTCTCAGAGACTCTGTCGCGCCCTTGATACACGCTTGAAGAGGGCTCTGAATGAGCCAC	40862
QY	1542	TTCTGCGCTCTGTACGTTGAGAAGCAAAAGAGTCCATGAGCTTTAAACAGCAGCTT	1601
Db	40861	TTCTGCGCTCTGTACGTTGAGAAGCAAAAGAGTCCATGAGCTTTAAACAGCAGCTT	40802
QY	1602	CCATATCACTCCGCTTTGATGTAAGGGTTCCTGTTTGCCCTCTGAGGAAAGAGTAG	1661
Db	40801	CCATATCACTCCGCTTTGATGTAAGGGTTCCTGTTTGCCCTCTGAGGAAAGAGTAG	40742
QY	1662	GAGGCCACTGAGAGCTCTGCTGAGGCTGTCCCGTTTCCCTGCGGGGTGAACAGAGCTTCT	1721
Db	40741	GAGGCCACTGAGAGCTCTGCTGAGGCTGTCCCGTTTCCCTGCGGGGTGAACAGAGCTTCT	40682
QY	1722	GCAATGGGTCTCTCTGTGGGTGACGAGCCTTAATGCAACACCCCAAGAGACACCCCTGCA	1781
Db	40681	GCAATGGGTCTCTCTGTGGGTGACGAGCCTTAATGCAACACCCCAAGAGACACCCCTGCA	40622
QY	1782	CGCCCTTCACACTGTCTTTCGAGACTCAAGACAAAGATTGTTCCGTTGGCATTAACAG	1841
Db	40621	CGCCCTTCACACTGTCTTTCGAGACTCAAGACAAAGATTGTTCCGTTGGCATTAACAG	40562
QY	1842	CTTCCAGAAGCTGCGCTTCGATTAACAGAACTCGACTTGATGCAATCTTCTTCG	1901
Db	40561	CTTCCAGAAGCTGCGCTTCGATTAACAGAACTCGACTTGATGCAATCTTCTTCG	40502
QY	1902	CTTCCAGACACTGTCCGATTTGCGGAAAAATTCGGGTGATGTCAAGGAGATCTTCCAG	1961
Db	40501	CTTCCAGACACTGTCCGATTTGCGGAAAAATTCGGGTGATGTCAAGGAGATCTTCCAG	40442
QY	1962	AGATGAGTCCGCTGAGGCATGTCTGTGTCCTCTATAG 2001	
Db	40441	AGATGAGTCCGCTGAGGCATGTCTGTGTCCTCTATAG 40402	

RESULT	12
AC024580 LOCUS	AC024580      193609 bp    DNA     linear   PRI 21-DEC-2001
DEFINITION	Homo sapiens chromosome 19 clone CTD-2621I17, complete sequence.
ACCESSION	AC024580
VERSION	AC024580.6   GI:17975240
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  1 (bases 1 to 193609) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission  Unpublished 2 (bases 1 to 193609) DOE Joint Genome Institute. Direct Submission Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 3 (bases 1 to 193609) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Submitted (13-NOV-2001) DOE Joint Genome Inetitue, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 4 (bases 1 to 193609) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Submitted (21-DEC-2001) DOE Joint Genome Inetitue, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Dec 21, 2001 this sequence version replaced gi:16905144. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.bngc.stanford.edu Quality: Phrap Quality >=40 99.4% of Sequence;



FEATURES		Estimated Total Number of Errors is 0.9.	
source		Location/Qualifiers	
		1. 193609	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/chromosome="19"	
		/clone="CTD-2621117"	
ORIGIN			
Query Match		38.4%;	Score 1549; DB 9; Length 193609;
Best Local Similarity		99.9%;	Pred. No. 0;
Matches 1599; Conservative		0; Mismatches	1; Indels 0; Gaps 0;
QY	402	AGGACATGAGGATGACATATGAGATCAAGAGTCACTGATGATCAACCAATTCGCTGAGGA	461
DB	22257	AGGACATGAGGATGACATATGAGATCAAGAGTCACTGATGATCAACCAATTCGCTGAGGA	22316
QY	462	GGAGATGTACGTCGATGTTTGAATAACACTGCTGCTGATCGGCGGGAATGCAACGTT	521
DB	22317	GGAGATGTACGTCGATGTTTGAATAACACTGCTGCTGATCGGCGGGAATGCAACGTT	22376
QY	522	GGCTGCTGCTTTGATTGATTCAGACCGATGCGGCTTCGCGCTCGACCGATGATTCGACCG	581
DB	22377	GGCTGCTGCTTTGATTGATTCAGACCGATGCGGCTTCGCGCTCGACCGATGATTCGACCG	22436
QY	582	AAAGTCAGGAATTTGGGAATTCGCTTACCCAGAGAGATCTGCTGCTGCGCGCAAG	641
DB	22437	AAAGTCAGGAATTTGGGAATTCGCTTACCCAGAGAGATCTGCTGCTGCGCGCAAG	22496
QY	642	TGAGCTCTACAGGGAATGTTCTCTACGCTTCTTCCCTCCCGTTAGAGATGAGAGG	701
DB	22497	TGAGCTCTACAGGGAATGTTCTCTACGCTTCTTCCCTCCCGTTAGAGATGAGAGG	22556
QY	702	GAGAAGAGAGAGATGTCAGAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCC	761
DB	22557	GAGAAGAGAGAGATGTCAGAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCC	22616
QY	762	GGTGAAGAGATCATATGCTCCGACAGAAAGGCTGTTGTTTCAATTCATGACGTTTCATGA	821
DB	22617	GGTGAAGAGATCATATGCTCCGACAGAAAGGCTGTTGTTTCAATTCATGACGTTTCATGA	22676
QY	822	CCGAGGCTCTGCTCTCAAGATGACCAAAAGCTCTGCAAAAGCTGCGGCTGAGAGAGCC	881
DB	22677	CCGAGGCTCTGCTCTCAAGATGACCAAAAGCTCTGCAAAAGCTGCGGCTGAGAGAGCC	22736
QY	882	TCCGTTCACTCTATAGCAGATCTGCTGAGAAAGTCTGCTCTGATCTCTTCTGAT	941
DB	22737	TCCGTTCACTCTATAGCAGATCTGCTGAGAAAGTCTGCTCTGATCTCTTCTGAT	22796
QY	942	CGTCACTGACAGAGCTGCGGCAAGAGAGCTCAAGTCAAGAGTCTGCTCTCCCGTTA	1001
DB	22797	CGTCACTGACAGAGCTGCGGCAAGAGAGCTCAAGTCAAGAGTCTGCTCTCCCGTTA	22856
QY	1002	CCCTGTTAGTAAAGAAATTCGCGGGAACAAAGAAATCCATCTGCTCTTGAAGCGGGAT	1061
DB	22857	CCCTGTTAGTAAAGAAATTCGCGGGAACAAAGAAATCCATCTGCTCTTGAAGCGGGAT	22916
QY	1062	TGATGAGCATCAGAAAGACAAAGGTTGGTGCATCATGAACAACCGTGAAGTGTCTCA	1121
DB	22917	TGATGAGCATCAGAAAGACAAAGGTTGGTGCATCATGAACAACCGTGAAGTGTCTCA	22976
QY	1122	CCAGTGCAGGTCGCCCGCTGCGGCTCTCTCATCTGCGTGGCCCTGACGCTCAAGACGT	1181
DB	22977	CCAGTGCAGGTCGCCCGCTGCGGCTCTCTCATCTGCGTGGCCCTGACGCTCAAGACGT	23036
QY	1182	GGTGGGGGAGAGCTGCGCCCTTCAACAAAGCTCAAGGCTGCAAGCGCGCTTTTGT	1241
DB	23037	GGTGGGGGAGAGCTGCGCCCTTCAACAAAGCTCAAGGCTGCAAGCGCGCTTTTGT	23096
QY	1242	GTTTCATCAGTCAACCCCTCGAGGCGTGGTCCGCGCTGCTCAATCTTGAAGAAAGAT	1301
DB	23097	GTTTCATCAGTCAACCCCTCGAGGCGTGGTCCGCGCTGCTCAATCTTGAAGAAAGAT	23156

QY	1302	TGTCCTGAAGGCTTCTGCGGATGATGCTGAGAGGAGTGTGAATPAGAACTCATGTT	1361
DB	23157	TGTCCTGAAGGCTTCTGCGGATGATGCTGAGAGGAGTGTGAATPAGAACTCATGTT	23216
QY	1362	TGATGATGACGACCTCATATGATTCAGAGACTCGGAGAGTCTGAGCTCCGCTCTGTTCA	1421
DB	23217	TGATGATGACGACCTCATATGATTCAGAGACTCGGAGAGTCTGAGCTCCGCTCTGTTCA	23276
QY	1422	CATGAACATCTTCTCCAGACAGCCATGTGAGAGTACTAACCTTCTTCCACTCAG	1481
DB	23277	CATGAACATCTTCTCCAGACAGCCATGTGAGAGTACTAACCTTCTTCCACTCAG	23336
QY	1482	TCTCCAGACTCTGTCGCGCTGTTGATCACTGATGAGAGGCTGGAATGAGCCAGC	1541
DB	23337	TCTCCAGACTCTGTCGCGCTGTTGATCACTGATGAGAGGCTGGAATGAGCCAGC	23396
QY	1542	TCTCTGCCCTCTGATGATGAGAAACAAAGAGTCCATGAGCTTAAACAGGAGCTT	1601
DB	23397	TCTCTGCCCTCTGATGATGAGAAACAAAGAGTCCATGAGCTTAAACAGGAGCTT	23456
QY	1602	CGATATCCACTGCTTGGATGATGAAGCTTCTTGTGCTGCTGAGCGAAGCTAAG	1661
DB	23457	CGATATCCACTGCTTGGATGATGAAGCTTCTTGTGCTGCTGAGCGAAGCTAAG	23516
QY	1662	GAGGCACTGAGAGTCTGCTGCGCTGCTCCGCTCCGCGGAGTGAAGCAGACCTCT	1721
DB	23517	GAGGCACTGAGAGTCTGCTGCGCTGCTCCGCTCCGCGGAGTGAAGCAGACCTCT	23576
QY	1722	GCATGAGGCTCTCTGTTGGGCTCAGCAGCTTAATGCAACCCAGAGAGACCCCTGGA	1781
DB	23577	GCATGAGGCTCTCTGTTGGGCTCAGCAGCTTAATGCAACCCAGAGAGACCCCTGGA	23636
QY	1782	CGCCTTCACTGCTCTTTTGAAGCTCAAGACAAAGATTTGTCGCTTGGATTAACAG	1841
DB	23637	CGCCTTCACTGCTCTTTTGAAGCTCAAGACAAAGATTTGTCGCTTGGATTAACAG	23696
QY	1842	CTTCCAGAAAGTGTGGCTTCGATTAACCAAGACTGAGCTTATGATCTTCTCTG	1901
DB	23697	CTTCCAGAAAGTGTGGCTTCGATTAACCAAGACTGAGCTTATGATCTTCTCTG	23756
QY	1902	CCTCCAGACGTCGCTATTTGCGGAAATTCGCGGATGATCAAGGATCTTCCCAAG	1961
DB	23757	CCTCCAGACGTCGCTATTTGCGGAAATTCGCGGATGATCAAGGATCTTCCCAAG	23816
QY	1962	AGATGAGTCCGCTGAGGATGCTCTGATGCTCTGATG	2001
DB	23817	AGATGAGTCCGCTGAGGATGCTCTGATGCTCTGATG	23856

RESULT 13	
AX427588	1157 bp DNA linear PAT 20-JUN-2002
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
source	
ORIGIN	

Query Match 27.3%; Score 1102; DB 6; Length 1157;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 757 GCTCCGGTACGAGATCATGTCCCGACCAAGAGGCTGTGTTTCATCATTTGACGTTTC 816  
 DB 4 GCTCCGGTACGAGATCATGTCCCGACCAAGAGGCTGTGTTTCATCATTTGACGTTTC 63

QY 817 GATGACCTGGGCTGTCTCTCAACAATGACCAAGAGCTGTGCAAGAGCTGGGCTGAGAG 876  
 DB 64 GATGACCTGGGCTGTCTCTCTCAACAATGACCAAGAGCTGTGCAAGAGCTGGGCTGAGAG 123

QY 877 CAGCTCCGTTACCTCTATACGAGTCTGTGAGAGAGTCTGTCTCTGAGTCTTC 936  
 DB 124 CAGCTCCGTTACCTCTATACGAGTCTGTGAGAGAGTCTGTCTCTGAGTCTTC 183

QY 937 CTGATCGTACCTGTACAGAGCTGGGCAAGAGAGTCTGAGAGTGTGTCTTC 996  
 DB 184 CTGATCGTACCTGTGTACAGAGCTGGGCAAGAGAGTCTGAGAGTGTGTCTTC 243

QY 997 CGTTACCTGTAGTATAGAGGAATCTCCGGGGAACAAATTCATCTGCTCTGAGGCG 1056  
 DB 244 CGTTACCTGTATAGTATAGGAATCTCCGGGGAACAAATTCATCTGCTCTGAGGCG 303

QY 1057 GGGATTGGTGAACATCAGAAAGACAGAGGTTGCGTGCAATCATGAAACCGTGAAGCTG 1116  
 DB 304 GGGATTGGTGAACATCAGAAAGACAGAGGTTGCGTGCAATCATGAAACCGTGAAGCTG 363

QY 1117 CTGACACAGTGCAGAGTCCCGCGCTGCTCTCATCTGTGCTGCTCTGAGTCTGAG 1176  
 DB 364 CTGACACAGTGCAGAGTCCCGCGCTGCTCTCATCTGTGCTGCTCTGAGTCTGAG 423

QY 1177 GACGCGTGGGGGAGAGGCTGCGCCCTTCAACCAAGCTCAAGGCTGACGCGCT 1236  
 DB 424 GACGCGTGGGGGAGAGGCTGCGCCCTTCAACCAAGCTCAAGGCTGACGCGCT 483

QY 1237 TTTGTGTTTCATCAGCTCACTCCCTGAGAGCGTGTCTGCGCTGTCTCATCTGAGAGAA 1296  
 DB 484 TTTGTGTTTCATCAGCTCACTCCCTGAGAGCGTGTCTGCGCTGTCTCATCTGAGAGAA 543

QY 1297 AGAGTTGTCCTGAAGCGCTTCTGCGGTATGCTGTGAGAGGAGTGTGAATAGAGTCA 1356  
 DB 544 AGAGTTGTCCTGAAGCGCTTCTGCGGTATGCTGTGAGAGGAGTGTGAATAGAGTCA 603

QY 1357 GTGTTGATGAGAGAGCTCATGTTGAGAGAGCTGGGGAGTCTGAGCTCCGCTCG 1416  
 DB 604 GTGTTGATGAGAGAGCTCATGTTGAGAGAGCTGGGGAGTCTGAGCTCCGCTCG 663

QY 1417 TTTCAATGAACATCTCTCTCCACAGACCACTGTGAGAGTACTAGACCTTCTTCAC 1476  
 DB 664 TTTCAATGAACATCTCTCTCCACAGACCACTGTGAGAGTACTAGACCTTCTTCAC 723

QY 1477 CTGAGTCTCAGAGATCTCTGTGCGCTGTGTAAGTGTGAGAGGCTGTGAATTCAG 1536  
 DB 724 CTGAGTCTCAGAGATCTCTGTGCGCTGTGTAAGTGTGAGAGGCTGTGAATTCAG 783

QY 1537 CCAGCTCTCTGCGCTCTGAGAGTGAAGAGCAAGAGAGTCCAGAGCTTAAACAGCA 1596  
 DB 784 CCAGCTCTCTGCGCTCTGAGAGTGAAGAGCAAGAGAGTCCAGAGCTTAAACAGCA 843

QY 1597 GAGCTTCATATCACTCTCTGAGTGAAGAGCTTCTGTGTTGAGCTGTGAGAGCAAGC 1656  
 DB 844 GAGCTTCATATCACTCTCTGAGTGAAGAGCTTCTGTGTTGAGCTGTGAGAGCAAGC 903

QY 1657 GTAAGAGGCACTGAGAGTCTGTGAGAGTCTGTGAGAGTCTGTGAGAGTCTGTGAG 1716  
 DB 904 GTAAGAGGCACTGAGAGTCTGTGAGAGTCTGTGAGAGTCTGTGAGAGTCTGTGAG 963

QY 1717 CTTCGCACTGGGCTCTCTGTGTTGGTGAAGAGCTTATGCAACCAAGAGAGAGC 1776  
 DB 964 CTTCGCACTGGGCTCTCTGTGTTGGTGAAGAGCTTATGCAACCAAGAGAGAGC 1023

QY 1777 CTGAGAGCTTCACTGTCTTTTCAGAGACTCAAGAGAGTGTGTTGCTGTGCAATTA 1836

DB 1024 CTGAGAGCTTCCACTGTCTTTTCAGAGCTCAAGAGAGTGTGTTGCTTGGCATTA 1083

QY 1837 AACAGCTTCCAGAGAGTGTGCTTCCAGATTAAACAGAGAGCTGTGATGATCTTCC 1896

DB 1084 AACAGCTTCCAGAGAGTGTGCTTCCAGATTAAACAGAGAGCTGTGATGATCTTCC 1143

QY 1897 TTTGCTCTCAGC 1909

DB 1144 TTTGCTCTCAGC 1156

RESULT 14  
 AC012107/c  
 LOCUS  
 DEFINITION Homo sapiens clone Rpl1-45K21, WORKING DRAFT SEQUENCE, 25 unordered  
 pieces.  
 ACCESSION AC012107  
 VERSION AC012107.2 GI:7329252  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 167509)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens, clone Rpl1-45K21  
 Unpublished  
 2 (bases 1 to 167509)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., DeRellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, D., Gardyna, S., Grant, G., Hagoe, B., Harford, A., Horton, L.,  
 Howland, D., C., Johnson, R., Jones, C., Kahn, L., Karate, A., Klein, J.,  
 Lehoczek, J., Lien, C., Locke, K., Macdonald, P., Marquis, A.,  
 McMan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, D.,  
 Teste, S., Tittell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 26, 2000 this sequence version replaced gi:6088020.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L1180  
 Center clone name: 45\_K21  
 ----- Summary Statistics  
 Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator-amersham; 4% of reads  
 Chemistry: Dye-terminator Big Dye; 96% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 145437 bases at least Q40  
 Consensus quality: 155496 bases at least Q30  
 Consensus quality: 159832 bases at least Q20  
 Insert size: 176000; agarose-fp  
 Insert size: 165109; sum-of-contigs  
 Quality coverage: 3.8 in Q20 bases; agarose-fp  
 Quality coverage: 4.1 in Q20 base.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 25 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is



\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1061: contig of 1061 bp in length  
1062 1161: gap of 100 bp  
1162 2381: contig of 1220 bp in length  
2382 2481: gap of 100 bp  
2482 3771: contig of 1250 bp in length  
3772 3871: gap of 100 bp  
3872 4949: contig of 1078 bp in length  
4950 5049: gap of 100 bp  
5050 6119: contig of 1070 bp in length  
6120 6220: gap of 100 bp  
6220 8023: contig of 1804 bp in length  
8024 8123: gap of 100 bp  
8124 10555: contig of 2432 bp in length  
10556 12768: contig of 2113 bp in length  
12769 12868: gap of 100 bp  
12869 15722: contig of 2854 bp in length  
15723 15822: gap of 100 bp  
15823 17959: contig of 2137 bp in length  
17960 18059: gap of 100 bp  
18060 21771: contig of 3712 bp in length  
21772 21871: gap of 100 bp  
21872 24539: contig of 2668 bp in length  
24540 24639: gap of 100 bp  
24640 28144: contig of 3505 bp in length  
28145 28244: gap of 100 bp  
28245 31503: contig of 3259 bp in length  
31504 31603: gap of 100 bp  
31604 35858: contig of 4255 bp in length  
35859 35958: gap of 100 bp  
35959 39610: contig of 3652 bp in length  
39611 39710: gap of 100 bp  
39711 42759: contig of 3049 bp in length  
42760 42859: gap of 100 bp  
42860 48014: contig of 5155 bp in length  
48015 48114: gap of 100 bp  
48115 52263: contig of 4149 bp in length  
52264 52363: gap of 100 bp  
52364 56374: contig of 4011 bp in length  
56375 56474: gap of 100 bp  
56475 63102: contig of 6628 bp in length  
63103 72687: contig of 9485 bp in length  
72688 72787: gap of 100 bp  
72788 97799: contig of 25012 bp in length  
97800 97899: gap of 100 bp  
97900 134325: contig of 36426 bp in length  
134326 167509: contig of 33084 bp in length.  
134426

## FEATURES

## SOURCE

1.167509  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-45K21"  
/clone\_lib="RP11-45K21 Human Male BAC"  
1.1061  
/note="assembly\_fragment"  
1162.2381  
/note="assembly\_fragment"  
2482.3771  
/note="assembly\_fragment"  
3872.4949  
/note="assembly\_fragment"  
5050.6119  
/note="assembly\_fragment"  
6220.8023  
/note="assembly\_fragment"

misc\_feature 8124..10555  
/note="assembly\_fragment"  
misc\_feature 10556..12768  
/note="assembly\_fragment"  
misc\_feature 12869..15722  
/note="assembly\_fragment"  
misc\_feature 15823..17959  
/note="assembly\_fragment"  
misc\_feature 18060..21771  
/note="assembly\_fragment"  
misc\_feature 21872..24539  
/note="assembly\_fragment"  
misc\_feature 24640..28144  
/note="assembly\_fragment"  
misc\_feature 28245..31503  
/note="assembly\_fragment"  
misc\_feature 31604..35858  
/note="assembly\_fragment"  
misc\_feature 35959..39610  
/note="assembly\_fragment"  
misc\_feature 39711..42759  
/note="assembly\_fragment"  
misc\_feature 42860..48014  
/note="assembly\_fragment"  
misc\_feature 48115..52263  
/note="assembly\_fragment"  
misc\_feature 52364..56374  
/note="assembly\_fragment"  
misc\_feature 56475..63102  
/note="assembly\_fragment"  
misc\_feature 63203..72687  
/note="assembly\_fragment"  
misc\_feature 72788..97799  
/note="assembly\_fragment"  
misc\_feature 97900..134325  
/note="assembly\_fragment"  
misc\_feature 134426..167509  
/note="assembly\_fragment"  
misc\_feature clone\_end:SP6  
vector\_side:left"

## ORIGIN

Query Match 22.3%; Score 900; DB 2; Length 167509;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 900; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1102 AACAAACGCTGAGCTGCTGACACAGCCAGGTCGCCGCGGCTCTCTCATCTGCCTG 1161  
DB 12664 AACAAACGCTGAGCTGCTGACACAGCCAGGTCGCCGCGGCTCTCTCATCTGCCTG 12605  
QY 1162 GCCCTGACGCTGACGAGCAGTGGTGGGAGAGCGTCGCCCTTCAACCAACGCTGACA 1221  
DB 12604 GCCCTGACGCTGACGAGCAGTGGTGGGAGAGCGTCGCCCTTCAACCAACGCTGACA 12555  
QY 1222 GGCCTGACGCGCCTTTTGTGTTTCATAGCTCAACCCCTGAGGCGTCCGGCGCTGT 1281  
DB 12544 GGCCTGACGCGCCTTTTGTGTTTCATAGCTCAACCCCTGAGGCGTCCGGCGCTGT 12485  
QY 1282 CTCATCTGAGAGAAAGTGTCTGAAAGGCTTTCGCCGTATGCTGTGAGAGGATG 1341  
DB 12484 CTCATCTGAGAGAAAGTGTCTGAAAGGCTTTCGCCGTATGCTGTGAGAGGATG 12425  
QY 1342 TGGAAATAGAGTCAAGTGTGATGATGATGACAGACCTCATGTTTCAAGAGCTCGGGAGTCT 1401  
DB 12424 TGGAAATAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 12365  
QY 1402 GAGCTCCGCTGCTGTGTTTACATGAATCCCTTCTCCGACACAGCCACTGTGAGAGTAC 1461  
DB 12364 GAGCTCCGCTGCTGTGTTTACATGAATCCCTTCTCCGACACAGCCACTGTGAGAGTAC 12305  
QY 1462 TACACCTTCTCCACCTCAAGTCTCAAGACCTTCTGTCGCCGCTGTACTACTGTAGAG 1521



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 17:31:03 ; Search time 2009 Seconds  
(without alignments)  
11889.575 Million cell updates/sec

Title: US-10-066-521-5  
Perfect score: 4035

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 beqs, 2959870667 residues

**word size :**

Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

### Post-processing: Listing first 45 summaries

```
Database :
1: geneeqgn16bdc04.*
2: geneeqgn19908.*
3: geneeqgn20008.*
4: geneeqgn2001as.*
5: geneeqgn2001bs.*
6: geneeqgn2002as.*
7: geneeqgn2002bs.*
8: geneeqgn2003as.*
9: geneeqgn2003bs.*
10: geneeqgn2003cs.*
11: geneeqgn2003ds.*
12: geneeqgn2004as.*
13: geneeqgn2004bs.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4035	100.0	4035	6	AAL44356	Aa1444356 Human PYR
2	4035	100.0	4035	10	ADP944744	AdP194744 Human PYR
3	2917	72.3	3489	6	AAD41224	Aad41224 Human EMB
4	2917	72.3	3926	8	ADA45218	Ada45218 Human MAT
5	2862	70.9	3830	8	ADA45220	Ada45220 Human MAT
6	2863	69.5	5859	6	AAL47135	Aa147135 Human MAT
7	2803	69.5	6939	6	AAL47131	Aa147131 Pyrin dom
8	2803	69.5	6939	6	AAL47140	Aa147140 Pyrin dom
9	2801	69.4	3226	6	ABK97181	ABK97181 Human NOV
10	2713	67.2	3900	6	ABK48628	ABK48628 Human MAT
11	2713	67.2	3900	6	AAD49018	Aad49018 Human MAT
12	1102	27.3	1157	6	ABK48609	ABK48609 Human MAT
13	1102	27.3	1157	8	AAD48999	Aad48999 Human MAT
14	657	16.3	1075	6	ABK48610	ABK48610 Human MAT
15	657	16.3	1075	8	AAD49000	Aad49000 Human MAT
16	316	7.8	2820	12	ADM62078	Adm62078 Human CDH
17	24	0.6	24	6	ABK48618	ABK48618 GSP #1 us
18	24	0.6	24	6	ABK48619	ABK48619 GSP #2 us
19	24	0.6	24	6	ABK97577	ABK97577 Human NOV
20	24	0.6	24	8	AAD49009	Aad49009 Human MAT

C	21	24	0.6	24	8	AAD49008	Human	MAT	Aad49008
C	22	24	0.6	24	12	ADN62477	Human	NOV	Adn62477
C	23	22	0.5	22	6	ABK48612	Human	MAT	Abk48612
C	24	22	0.5	22	6	AAD49002	Human	MAT	Aad49002
C	25	22	0.5	26	6	ABK48614	Human	MAT	Abk48614
C	26	22	0.5	26	8	AAD49004	Human	MAT	Aad49004
C	27	22	0.5	27	6	ABK48622	5'-primer		Abk48622
C	28	22	0.5	27	8	AAD49012	Human	MAT	Aad49012
C	29	22	0.5	3447	6	ABK48611	Mouse	MAT	Abk48611
C	30	22	0.5	3447	8	AAD49001	Mouse	MAT	Aad49001
C	31	22	0.5	9516	4	AAK81394	Human	imm	Aak81394
C	32	22	0.5	9516	4	AAK81394	Human	imm	Aak81394
C	33	22	0.5	13192	4	AAK73555	Human	imm	Aak73555
C	34	22	0.5	13192	4	AAK73556	Human	imm	Aak73556
C	35	22	0.5			AAK81396	Human	imm	Aak81396
C	36	21	0.5	21	6	ABK48613	Human	MAT	Abk48613
C	37	21	0.5	21	6	ABK48624	Human	MAT	Abk48624
C	38	21	0.5	21	6	ABK48625	Human	MAT	Abk48625
C	39	21	0.5	21	6	ABX97578	Human	NOV	Abx97578
C	40	21	0.5	21	6	ABX97576	Human	NOV	Abx97576
C	41	21	0.5	21	8	ADA45224	Human	MAT	Ada45224
C	42	21	0.5	21	8	ADA45222	Human	MAT	Ada45222
C	43	21	0.5	21	8	AAD49014	Human	MAT	Aad49014
C	44	21	0.5	21	8	AAD49015	Human	MAT	Aad49015
C	45	21	0.5	21	8	AAD49003	Human	MAT	Aad49003

## ALIGNMENTS

RESULT 1	
AL44356	AL44356 standard; cDNA; 4035 BP.
ID	AL44356
XX	
AC	AL44356;
XX	
D7	31-OCT-2002 (first entry)
XX	
DE	Human PYRIN-5 cDNA sequence.
XX	
KW	Human; gene; 89; gene therapy; PYRIN; stress-related response;
KW	apoptotic response; inflammatory response; inflammatory disorder;
KW	immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW	leukemia; autoimmune disorder; arthritis; neurological disease;
KW	Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW	tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW	transcript profiling; PYRIN-5.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..4035
FT	/*tag= a
FT	/product= "Human PYRIN-5"
XX	
PN	MO200261049-A2.
XX	
PD	08-AUG-2002.
XX	
PE	31-JAN-2002; 2002MO-US002967.
XX	
PR	31-JAN-2001; 2001US-0265231P.
PR	10-SEP-2001; 2001US-0318645P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
PA	(AMHP) WYETH.
XX	
PI	Bertin J, Wang W, Blatcher M;
XX	
DR	WPI: 2002-627477/67.
DR	P-PSDB; AAO15585.
XX	
XX	New PYRIN polypeptides and nucleic acids useful for modulating and



Db 1741 GGTCAAGACCTAATGCCACACCCAGAGAGACCTTGAGCGCTTCCACTGTCTTTC 1800  
Qy 1801 GAGATTCAGAACAAAGATTGTTGGCTTGCGATTAAACAGTTTCCAAAGATGTCCTT 1860  
Db 1801 GAGACTCAAGAACAAAGATTGTTGGCTTGCGATTAAACAGTTTCCAAAGATGTCCTT 1860  
Qy 1861 CCGATTAAACAGAACCTTGAGATTGATCTTCTGCTGCTGAGAGCTGTCCTGAT 1920  
Db 1861 CCGATTAAACAGAACCTTGAGATTGATCTTCTGCTGCTGAGAGCTGTCCTGAT 1920  
Qy 1921 TTGCGAAAAATTCCGGGTGAGATGTCAAAGGATCTTCCAAAGATGATGTCGCTGAGCA 1980  
Db 1921 TTGCGAAAAATTCCGGGTGAGATGTCAAAGGATCTTCCAAAGATGATGTCGCTGAGCA 1980  
Qy 1981 TGTCTGTGTGTCCTCTATGATGTCGGGATTAAGACCTTCAATTGAGAGAGAGTGGAAAT 2040  
Db 1981 TGTCTGTGTGTCCTCTATGATGTCGGGATTAAGACCTTCAATTGAGAGAGAGTGGAAAT 2040  
Qy 2041 TTCTGCTCAATGCTTGGGACCCACACCTGCGGAGCTGAGCTGCGGAGAGCAATC 2100  
Db 2041 TTCTGCTCAATGCTTGGGACCCACACCTGCGGAGCTGAGCTGCGGAGAGCAATC 2100  
Qy 2101 CTGACAGAGCGGGCCATGAGAACCTGTGTGTCAGAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGGCCATGAGAACCTGTGTGTCAGAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTTGAAGATGACAGATTAACCTGTGTGTCAGACCTCTGAGAAATC 2220  
Db 2161 CAGACCTGATGTTTGAAGATGACAGATTAACCTGTGTGTCAGACCTCTGAGAAATC 2220  
Qy 2221 GTCAATGGCAACCGTAACCTTAAGATCCCTCAACTTGGGAGGACCACTGTAAGGAAG 2280  
Db 2221 GTCAATGGCAACCGTAACCTTAAGATCCCTCAACTTGGGAGGACCACTGTAAGGAAG 2280  
Qy 2281 GATGTAAGGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGGAGTCTTGAAG 2340  
Db 2281 GATGTAAGGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGGAGTCTTGAAG 2340  
Qy 2341 CTGATTCCTGTGATGATGACCCATGCTGTAACCTGAAGATCTCCAAATCTTTCAGAC 2400  
Db 2341 CTGATTCCTGTGATGATGACCCATGCTGTAACCTGAAGATCTCCAAATCTTTCAGAC 2400  
Qy 2401 TCCCCAGCGTAATCTCTGAGCTGCGAGAGAACAGATGACAGACAGAGGAATAATG 2460  
Db 2401 TCCCCAGCGTAATCTCTGAGCTGCGAGAGAACAGATGACAGACAGAGGAATAATG 2460  
Qy 2461 CCTCTCAGTGAATGCTTGAAGATCTCCAGTGCCTGCGAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTGAATGCTTGAAGATCTCCAGTGCCTGCGAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGCAATCAAGCAAGGCTTGGAGATCTGAGCTTCAAGCTTGTGACAGACCGAGAC 2580  
Db 2521 TGTGCAATCAAGCAAGGCTTGGAGATCTGAGCTTCAAGCTTGTGACAGACCGAGAC 2580  
Qy 2581 TTGACACACCTGTGCTATCCAAACAGCCCTGGGAGAGAGATGTAATCTTACTGTGT 2640  
Db 2581 TTGACACACCTGTGCTATCCAAACAGCCCTGGGAGAGAGATGTAATCTTACTGTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGATGTCGAGAGGCTGATGCTGATCTGAGGACCTG 2700  
Db 2641 CGATCATGAGGCTTCCCACTGATGTCGAGAGGCTGATGCTGATCTGAGGACCTG 2700  
Qy 2701 GACAGGCTGAGCTGATGTTTCTTTCACCTTGGCTTAATGAGTCAATGCTGACGAC 2760  
Db 2701 GACAGGCTGAGCTGATGTTTCTTTCACCTTGGCTTAATGAGTCAATGCTGACGAC 2760  
Qy 2761 CTGAGCTTGAAGATGATGATGTCGAGAGATGTCGAGAGCTTGTGCGAGGTCAATG 2820  
Db 2761 CTGAGCTTGAAGATGATGATGTCGAGAGATGTCGAGAGCTTGTGCGAGGTCAATG 2820  
Qy 2821 AGAGAACCATCTTGTCAATCTCCAGAGACCTGAGATGTTGTAAGTGTCACTCCAGCG 2880  
Db 2821 AGAGAACCATCTTGTCAATCTCCAGAGACCTGAGATGTTGTAAGTGTCACTCCAGCG 2880

Qy 2881 TGTCTGAGAGATCTGCTGTGTGTGATCTCCAGAGAGAGACACTGAAAGGCTGATCTC 2940  
Db 2881 TGTCTGAGAGATCTGCTGTGTGTGATCTCCAGAGAGAGACACTGAAAGGCTGATCTC 2940  
Qy 2941 ACGGACATGCGCTGAGGATGAGCGTGGGATGCTGCTGCTGTCGAGGAGCTGAACAAAG 3000  
Db 2941 ACGGACATGCGCTGAGGATGAGCGTGGGATGCTGCTGCTGTCGAGGAGCTGAACAAAG 3000  
Qy 3001 AACAGTGTCTGAGAGACTCGGGTTGAAGGATGAGCTGATCTGATTTGCTGTGAG 3060  
Db 3001 AACAGTGTCTGAGAGACTCGGGTTGAAGGATGAGCTGATCTGATTTGCTGTGAG 3060  
Qy 3061 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
Db 3061 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
Qy 3121 AACTTCAGTCCCAAGAGATGATGAAGCTGTGTCGCTGCTGCTGCTGCTGCTGCTGCT 3180  
Db 3121 AACTTCAGTCCCAAGAGATGATGAAGCTGTGTCGCTGCTGCTGCTGCTGCTGCTGCT 3180  
Qy 3181 TTACAGATTAATGAGGCTGTGAGAAATGCGAGTAACTTGAAGATGAGAGCTGCTGAG 3240  
Db 3181 TTACAGATTAATGAGGCTGTGAGAAATGCGAGTAACTTGAAGATGAGAGCTGCTGAG 3240  
Qy 3241 GAAGTGCAGTACTCAAGCCCGAGTGTAAATGACGATGATGAGCTTCTTGTGATAA 3300  
Db 3241 GAAGTGCAGTACTCAAGCCCGAGTGTAAATGACGATGATGAGCTTCTTGTGATAA 3300  
Qy 3301 GATGACCGACCAAAATGAGACTTCTGCGGCTCCCTGAAGACCGGAGCATGCT 3360  
Db 3301 GATGACCGACCAAAATGAGACTTCTGCGGCTCCCTGAAGACCGGAGCATGCT 3360  
Qy 3361 GCTTGTGTGAGGAGATGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
Db 3361 GCTTGTGTGAGGAGATGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
Qy 3421 TTCAAGAGCAGTACAGATTTGCGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480  
Db 3421 TTCAAGAGCAGTACAGATTTGCGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480  
Qy 3481 CAGAGATGACAGAGTGGAGCAGAGCTCCCGAGACCCATGAGAGAGAGAGAGAGAGAG 3540  
Db 3481 CAGAGATGACAGAGTGGAGCAGAGCTCCCGAGACCCATGAGAGAGAGAGAGAGAGAG 3540  
Qy 3541 CAAGTAAATGTTGATGTTGATTAATCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600  
Db 3541 CAAGTAAATGTTGATGTTGATTAATCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600  
Qy 3601 GGGCTTGGATCAACAGTGTGATCAATGACAGAGGATGCTGCTGCTGCTGCTGCTGCTG 3660  
Db 3601 GGGCTTGGATCAACAGTGTGATCAATGACAGAGGATGCTGCTGCTGCTGCTGCTGCTG 3660  
Qy 3661 GAGCTGAGCTGAGAGGCTTGTGTCACAGTGTGATGACAGAGGATGCTGCTGCTGCTG 3720  
Db 3661 GAGCTGAGCTGAGAGGCTTGTGTCACAGTGTGATGACAGAGGATGCTGCTGCTGCTG 3720  
Qy 3721 CACTGAGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 3780  
Db 3721 CACTGAGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 3780  
Qy 3781 GTGTCTGTGATCACTGAGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 3840  
Db 3781 GTGTCTGTGATCACTGAGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 3840  
Qy 3841 CACAGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3900  
Db 3841 CACAGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3900  
Qy 3901 GCTGATGACACAGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3960  
Db 3901 GCTGATGACACAGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3960

OY		3961	TCCAACAGTGCCTGATGACCACACCGCGTGTCCTGTCTGACTCTGGAGCGCGCTGGCTCCAG	4021
Db		3961	TCCAACAGTGCCTGGAAGACACACGCGTGTCCTGTCTGACTCTGGAGCGCGCTGGCTCCAG	4021
OY		4021	GCGCTGCTGCTCTTA	4035
Db		4021	GCGCTGCTGCTCTTA	4035
RESULT 2				
ID	ADFP94744			
XX	ADFP94744 standard; cDNA; 4035 BP.			
XX				
AC	ADFP94744;			
XX				
DT	26-FEB-2004 (first entry)			
XX				
DE	Human PYRIN-5 coding sequence.			
XX				
KW	human; PYRIN; inflammatory disorder; inappropriate apoptosis;			
KW	inflammatory bowel disease; rheumatoid arthritis; diabetes;			
KW	multiple sclerosis; Grave's disease; contact dermatitis; psoriasis;			
KW	graft rejection; asthma; allergy; chronic obstructive pulmonary disease;			
KM	glomerulonephritis; infection; Alzheimer's disease; Parkinson's disease;			
KM	anemia; ischaemia; screening; chromosomal mapping; tissue typing;			
KW	forensic biology; pharmacogenomics; predictive medicine; gene; ss.			
XX				
OS	Homo sapiens.			
XX				
PN	WO2003089588-A2.			
PD	30-OCT-2003.			
XX				
PF	14-APR-2003; 2003WO-US011572.			
XX				
PR	17-APR-2002; 2002US-00124498.			
XX				
PA	(MILL-) MILLENNIUM PHARM INC.			
PA	(AMHP) WYTEH.			
Pt	Bertin J, Wang W, Blatcher M;			
DR	WPI: 2003-845527/78.			
Dr	P-P5DB; ADFP94745.			
PT	New nucleic acid molecules and polypeptides (e.g. PYRIN-2 or PYRIN-3)			
PT	useful for diagnosing, preventing or treating inflammation or disorders			
PT	associated with inappropriate apoptosis, in chromosomal mapping or in			
PT	pharmacogenomics.			
PS	Claim 4; SEQ ID NO 5; 19pp; English.			
XX				
CC	The invention comprises the amino acid and coding sequences of human			
CC	PYRIN proteins. The DNA and protein sequences of the invention are useful			
CC	in diagnosing, preventing and treating inflammatory disorders or			
CC	disorders associated with inappropriate apoptosis, such as: inflammatory			
CC	bowel disease, rheumatoid arthritis, diabetes, graft rejection, asthma,			
CC	Grave's disease, contact dermatitis, psoriasis, multiple sclerosis,			
CC	allergy, chronic obstructive pulmonary disease, glomerulonephritis,			
CC	infections, Alzheimer's disease, Parkinson's disease, anaemia and			
CC	ischaemia. The DNA and protein sequences of the invention may also be			
CC	used in screening assays, chromosomal mapping, tissue typing, forensic			
CC	biology, pharmacogenomics, predictive medicine, and in monitoring of			
CC	clinical trials. The present DNA sequence encodes a PYRIN protein of the			
CC	invention.			
XX				
SQ	Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;			
Query Match				
Best Local Similarity 100.0%; Score 4035; DB 10; Length 4035;				
Match 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	1	ATGGAAGAGACAATTCGCTCACCCTTTCCAGCTACGCGGCTGCAATGTCGTCTATGAG	60	

Db	1	ATGAGAGGAGCAAAATGCTACCTTTTTCAGCTACGGGCTGCATGGGTGCTCTATGAG	60
QY	61	CTAGACAAAGAAAGATTTTCAGACATTCAGAAATTTACTAAAGAAATCTTCAGAAATG	120
Db	61	CTAGACAAAGAAAGATTTTCAGACATTCAGAAATTTACTAAAGAAATCTTCAGAAATG	120
QY	121	ACCAATGCTCTATTTCCAAGTTTGAATCGAGATGCGCAACGTGGAATGTCTGGCACTC	180
Db	121	ACCAATGCTCTATTTCCAAGTTTGAATCGAGATGCGCAACGTGGAATGTCTGGCACTC	180
QY	181	CTCTTGAATGAGTATTTATGAGCATGCTGGGCTAGCTCATTTAGCATCTTTGAA	240
Db	181	CTCTTGAATGAGTATTTATGAGCATGCTGGGCTAGCTCATTTAGCATCTTTGAA	240
QY	241	AACATGAACTCGCAACCTCTCCGAGAAAGCAAGGATGACATGAAAAAAATTTTCAAA	300
Db	241	AACATGAACTCGCAACCTCTCCGAGAAAGCAAGGATGACATGAAAAAAATTTTCAAA	300
QY	301	GCTATGGAACAAGAGTGCCACAGCAGCAGAGACAGAGAACAGAAATTTTCAAACT	360
Db	301	GCTATGGAACAAGAGTGCCACAGCAGCAGAGACAGAGAACAGAAATTTTCAAACT	360
QY	361	ATGGAACAAGAGTGCCACAGCAGCAGAGACAGAGAACAGAAATTTTCAAACT	420
Db	361	ATGGAACAAGAGTGCCACAGCAGCAGAGACAGAGAACAGAAATTTTCAAACT	420
QY	421	TGGGACTATCAAGATCACTGATATACCAAAATTCCTGAGAGAGAGATGTACGTCTAGT	480
Db	421	TGGGACTATCAAGATCACTGATATACCAAAATTCCTGAGAGAGAGATGTACGTCTAGT	480
QY	481	TTTGAACAACTGCTGCTGACCTGGCCGGAATGCAAACTGTTGCTTTTGATTTCA	540
Db	481	TTTGAACAACTGCTGCTGACCTGGCCGGAATGCAAACTGTTGCTTTTGATTTCA	540
QY	541	GACCGGTGGGGCTTCCGGCTTCGACAGGTGTTTTCACAGGAAGTCAGAAATTTGGAAA	600
Db	541	GACCGGTGGGGCTTCCGGCTTCGACAGGTGTTTTCACAGGAAGTCAGAAATTTGGAAA	600
QY	601	TGGGCTCAGCAGAAAGGATGATGTCGTGCTGGGCGCAAGGTGAGCTCTACAGGGAATG	660
Db	601	TGGGCTCAGCAGAAAGGATGATGTCGTGCTGGGCGCAAGGTGAGCTCTACAGGGAATG	660
QY	661	TTCTCTACGTCTTCTCTCCCGTTTAAAGAGATGCAAGCGGAAGAAAGAGACAGTGC	720
Db	661	TTCTCTACGTCTTCTCTCCCGTTTAAAGAGATGCAAGCGGAAGAAAGAGACAGTGC	720
QY	721	ACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACGGAATCATGTCC	780
Db	721	ACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACGGAATCATGTCC	780
QY	781	CGACCAAGAAAGGCTGTTTCATCATGTAACGCTTCATGACCTGAGCTCTGTCTTAC	840
Db	781	CGACCAAGAAAGGCTGTTTCATCATGTAACGCTTCATGACCTGAGCTCTGTCTTAC	840
QY	841	AATGACAACAAAGCTCTGCAAAAGCTGGGCTGGAAGAGAGCTCCGTTACCCCTCATAGC	900
Db	841	AATGACAACAAAGCTCTGCAAAAGCTGGGCTGGAAGAGAGAGCTCCGTTACCCCTCATAGC	900
QY	901	AGTCTCTGAGGAAGTCTGCTCCCTGAGTCTTCTGATGTCACCGTCAGAGAGTGC	960
Db	901	AGTCTCTGAGGAAGTCTGCTCCCTGAGTCTTCTGATGTCACCGTCAGAGAGTGC	960
QY	961	GGCACAAGAAAGTCTCAAGTCAAGAGTGTGTCCTCCGTTACCTGTTAGTGAAGGATC	1020
Db	961	GGCACAAGAAAGTCTCAAGTCAAGAGTGTGTCCTCCGTTACCTGTTAGTGAAGGATC	1020
QY	1021	TCCGGGGAACAAAGATTCACCTTCTGAGGGCGGAGATTGGTGAACATCAAAAGACA	1080
Db	1021	TCCGGGGAACAAAGATTCACCTTCTGAGGGCGGAGATTGGTGAACATCAAAAGACA	1080
QY	1081	CAAGGTTGCTGCGATCATGAACAAACGTGAGCTGTCAGCAAGTCCAGGTGCCGCC	1140



1081 CAAGGGTTCGTCGATCATGAACAACCGTAGCTGCTGACCAAGTGCAGGTCGCCG 1140  
1141 GTGGGCTCTCTATCTGCTGGGCTCTGAGCTGCAAGAGTGTGGGGAGAGCTGCGC 1200  
1141 GTGGGCTCTCTATCTGCTGGGCTCTGAGCTGCAAGAGTGTGGGGAGAGCTGCGC 1200  
1201 CCTTCACCAACGCTCAAGAGCTGAGGCTGAGGCTGTTGTGTTTCAATGAGCTCAACCT 1260  
1201 CCTTCACCAACGCTCAAGAGCTGAGGCTGAGGCTGTTGTGTTTCAATGAGCTCAACCT 1260  
1261 CGAGGCTGGTCCGCGCTGTCTCAATCTGAGAGAAAGATTGTCTGAAGCGCTTTCG 1320  
1261 CGAGGCTGGTCCGCGCTGTCTCAATCTGAGAGAAAGATTGTCTGAAGCGCTTTCG 1320  
1321 CATTAGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTGTGATGTGATGCAAGCTCATG 1380  
1321 CATTAGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTGTGATGTGATGCAAGCTCATG 1380  
1321 CATTAGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTGTGATGTGATGCAAGCTCATG 1380  
1381 GTTCAAGAGCTCGGGAGTCTGAGCTCGTGTCTGTTCACATGAACATCTCTTCCCA 1440  
1381 GTTCAAGAGCTCGGGAGTCTGAGCTCGTGTCTGTTCACATGAACATCTCTTCCCA 1440  
1441 GACAGCACTGTGAGAGAGTCAACCTTCTTCACTCACTCACTCAAGAGCTTGTGTC 1500  
1441 GACAGCACTGTGAGAGAGTCAACCTTCTTCACTCACTCACTCAAGAGCTTGTGTC 1500  
1501 GCTTGTACTAGTGTGAGAGGCTGGAATGAGCCAGCTCTGTGCTCTGTACGTT 1560  
1501 GCTTGTACTAGTGTGAGAGGCTGGAATGAGCCAGCTCTGTGCTCTGTACGTT 1560  
1561 GAGAGACAAAGGCTCAATGAGCTTAAACAGGAGGCTTCAATTCACCTGCTTGG 1620  
1561 GAGAGACAAAGGCTCAATGAGCTTAAACAGGAGGCTTCAATTCACCTGCTTGG 1620  
1621 ATGAAGCGTTTCTGTTTGGCTCTGAGAGCAAGAGTGAAGAGGCACTGAGAGCTG 1680  
1621 ATGAAGCGTTTCTGTTTGGCTCTGAGAGCAAGAGTGAAGAGGCACTGAGAGCTG 1680  
1681 CTGGGCTGTCCCGTTCCCTGAGGAGTGAAGCAAGCTTCTGCACTGGGCTCTGTGTTG 1740  
1681 CTGGGCTGTCCCGTTCCCTGAGGAGTGAAGCAAGCTTCTGCACTGGGCTCTGTGTTG 1740  
1741 GGTGAGGCTCAATGAGCAACCCAGAGAGCAACCTGAGAGGCTTCACTGTCTTTC 1800  
1741 GGTGAGGCTCAATGAGCAACCCAGAGAGCAACCTGAGAGGCTTCACTGTCTTTC 1800  
1801 GAGACTCAAGCAAGAGTGTGCTGAGTGAATTAACAGCTTCAAGAGTGTGCTT 1860  
1801 GAGACTCAAGCAAGAGTGTGCTGAGTGAATTAACAGCTTCAAGAGTGTGCTT 1860  
1861 CGGATTAAACAGAACTTGAATGAGCTTCTGCTTCTGCTGAGCACTGTCCGAT 1920  
1861 CGGATTAAACAGAACTTGAATGAGCTTCTGCTTCTGCTGAGCACTGTCCGAT 1920  
1921 TTGGGGAATAATCGGGTGAATTCGAAGGATCTTCCGAAGAGTGTGCTGAGAG 1980  
1921 TTGGGGAATAATCGGGTGAATTCGAAGGATCTTCCGAAGAGTGTGCTGAGAG 1980  
1981 TGTCTGTGTGCTCTCTATGAGTGTGAGATTAAGCCCTCATTTGAGAGAGTGTGAGAT 2040  
1981 TGTCTGTGTGCTCTCTATGAGTGTGAGATTAAGCCCTCATTTGAGAGAGTGTGAGAT 2040  
2041 TTCTCTCCATCTTGGCAACCAACCACTGCGGAGCTGAGCTGGGAGAGCACTC 2100  
2041 TTCTCTCCATCTTGGCAACCAACCACTGCGGAGCTGAGCTGGGAGAGCACTC 2100  
2101 CTGAGAGAGGCGGCAATGAAGCCCTGTGTGCAAGCTGAGAGCTTCCCACTGCAAGATA 2160  
2101 CTGAGAGAGGCGGCAATGAAGCCCTGTGTGCAAGCTGAGAGCTTCCCACTGCAAGATA 2160  
2161 CAGAGCTGTGTGTAAGATGCAAGATTAACCTGCTGTGTGAGAGCACTCTGAGAGATC 2220  
2161 CAGAGCTGTGTGTAAGATGCAAGATTAACCTGCTGTGTGAGAGCACTCTGAGAGATC 2220

2221 GTTATGAGCAACCTTAACCTTAAGATCCCTCAACTTGGAGAGGCACTCACTGAAGAGAG 2280  
2221 GTTATGAGCAACCTTAACCTTAAGATCCCTCAACTTGGAGAGGCACTCACTGAAGAGAG 2280  
2281 GATGTAAAGATGAGCTGTGAAGCTTAAACCAACCAATGTTTGTGAGTCTTGAAG 2340  
2281 GATGTAAAGATGAGCTGTGAAGCTTAAACCAACCAATGTTTGTGAGTCTTGAAG 2340  
2341 CTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
2341 CTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
2401 TCCCCAGCTGGAATCTCTGAGCTGTGAGAGAAAGAGTGAAGAGCAAGAGAGTAAATG 2460  
2401 TCCCCAGCTGGAATCTCTGAGCTGTGAGAGAAAGAGTGAAGAGCAAGAGAGTAAATG 2460  
2461 CCTCTCAGTATGCTTGAAGAGTCTCCAGTGTGAGGCTGAGAGAGCTGATCTGAGAGAG 2520  
2461 CCTCTCAGTATGCTTGAAGAGTCTCCAGTGTGAGGCTGAGAGAGCTGATCTGAGAGAG 2520  
2521 TGTGATCAACAGCAAGGCTTGAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2580  
2521 TGTGATCAACAGCAAGGCTTGAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2580  
2581 TTGACACACTGTGCTTCAACCAACAGCTGTGAGAGAAAGAGTGAATTAATCTGTGT 2640  
2581 TTGACACACTGTGCTTCAACCAACAGCTGTGAGAGAAAGAGTGAATTAATCTGTGT 2640  
2641 GATCAGAGAGCTTCCCACTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2700  
2641 GATCAGAGAGCTTCCCACTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2700  
2701 GACAGGCTGTGAGTGTGTTTCTTTCACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2760  
2701 GACAGGCTGTGAGTGTGTTTCTTTCACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2760  
2761 CTGAGCTTGAAGTAAACCTGTGAGAGCAATGAGCTGGAAGCTTCTGTGAGAGTGTGAG 2820  
2761 CTGAGCTTGAAGTAAACCTGTGAGAGCAATGAGCTGGAAGCTTCTGTGAGAGTGTGAG 2820  
2821 AGAGAACATCTGATCTCAAGAGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2880  
2821 AGAGAACATCTGATCTCAAGAGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2880  
2881 TGTGTGAGAGTGTGCTGT 2940  
2881 TGTGTGAGAGTGTGCTGT 2940  
2941 ACGAGCAATGCTGT 3000  
2941 ACGAGCAATGCTGT 3000  
3001 AACAGTGTGTGAGCAAGCTGT 3060  
3001 AACAGTGTGTGAGCAAGCTGT 3060  
3061 GCACTCTCTGAGCTTCTGT 3120  
3061 GCACTCTCTGAGCTTCTGT 3120  
3121 AACTTCACTTCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180  
3121 AACTTCACTTCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180  
3181 TTACAGATTAATTTGGCTGTGAGAAATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240  
3181 TTACAGATTAATTTGGCTGTGAGAAATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240  
3241 GAGTGTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 3300  
3241 GAGTGTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 3300



OY		3301	GATGACCGACCAAAATTAGACTTATCTTCGGGCTCCGTGAAGAAGCCGGGACATGGCATTGT	3360
Db		3301	GATGACCGACCAAAATTAGACTTATCTTCGGGCTCCGTGAAGAAGCCGGGACATGGCATTGT	3360
OY		3361	GCGTTGCTGTGGGGGATGAACCACAAGACAGAAGACGCTGTGTGCCTTCTGGCTTGAGAC	3420
Db		3361	GCGTTGCTGTGGGGGATGAACCACAAGACAGAAGACGCTGTGTGCCTTCTGGCTTGAGAC	3420
OY		3421	TTCAAAGACAGTACACGATTTTGCAAAGTCTCTCTGCTGGCCACCGGCAAAATGTGAATGCC	3480
Db		3421	TTCAAAGACAGTACACGATTTTGCAAAGTCTCTCTGCTGGCCACCGGCAAAATGTGAATGCC	3480
OY		3481	CAGAGAGCTTGAACAAGTGGAGCAGAGGCTCCCCGCAACCCATGGACGGGACACGGAAACACAA	3540
Db		3481	CAGAGAGCTTGAACAAGTGGAGCAGAGGCTCCCCGCAACCCATGGACGGGACACGGAAACACAA	3540
OY		3541	CAAGATAAATAGTTAGTGTGGATTAATTCGGAGCTGAGTCTGAATACTGTAGAGCTCGAA	3600
Db		3541	CAAGATAAATAGTTAGTGTGGATTAATTCGGAGCTGAGTCTGAATACTGTAGAGCTCGAA	3600
OY		3601	GCGGCTTGGATCCAACAGTGTCTGATCATGAACAACGAGAGTATAGGCTGTGTCACTAGGAGAA	3660
Db		3601	GCGGCTTGGATCCAACAGTGTCTGATCATGAACAACGAGAGTATAGGCTGTGTCACTAGGAGAA	3660
OY		3661	GAGCTGAGACTGAGAGGGGCTTGTGTCCAACAGTGTGATGAACAACAGCGGCTGTCTTGGT	3720
Db		3661	GAGCTGAGACTGAGAGGGGCTTGTGTCCAACAGTGTGATGAACAACAGCGGCTGTCTTGGT	3720
OY		3721	CAGTGGAGCGGCTGGGCTTAGAGGGGCTGTGTCTTAAACAGTGTGATGAACAACAGCGGT	3780
Db		3721	CAGTGGAGCGGCTGGGCTTAGAGGGGCTGTGTCTTAAACAGTGTGATGAACAACAGCGGT	3780
OY		3781	GTGTCTCTGTGCACTGGGAGCGGCTGGGCTCGAAGGGCTTGTGTCCAACAGTGTCTGATGAC	3840
Db		3781	GTGTCTCTGTGCACTGGGAGCGGCTGGGCTCGAAGGGCTTGTGTCCAACAGTGTCTGATGAC	3840
OY		3841	CACAGCGGCTGTGGGCTGTGTGTCACTGGGAGCGGCTGGGCTTGTGTCCAACAGT	3900
Db		3841	CACAGCGGCTGTGGGCTGTGTGTCACTGGGAGCGGCTGGGCTTGTGTCCAACAGT	3900
OY		3901	GCTATATGACCAAGCGGTGTGTCTCGTGTCACTGGGAGGGGCTGGGCTGGAAGGGGCTTGTG	3960
Db		3901	GCTATATGACCAAGCGGTGTGTCTCGTGTCACTGGGAGGGGCTGGGCTGGAAGGGGCTTGTG	3960
OY		3961	TCCAACAATGTGTGATGAACAACAGCGGCTGTGTCTGTGCACTGGGAGCGGCTGGGCTCGAG	4020
Db		3961	TCCAACAATGTGTGATGAACAACAGCGGCTGTGTCTGTGCACTGGGAGCGGCTGGGCTCGAG	4020
OY		4021	GCGGCTGTGTCTTAA	4035
Db		4021	GCGGCTGTGTCTTAA	4035
RESULT 3				
AAD41224	ID	AAD41224	standard; cDNA; 3489 BP.	
XX	AC	AAD41224;		
XX	DT	30-OCT-2002	(first entry)	
XX	DE	Human EMBRY-1 cDNA.		
KW		Human; embryogenesis associated protein; AIDS; reproductive disorder;		
KW		infertility; endometriosis; endometrial tumour; inflammatory disorder;		
KW		autoimmune disorder; acquired immune deficiency syndrome; transgenic;		
KW		ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;		
OS		EMBRY-1; allergy; gene therapy; gene; ss.		
XX		Homo sapiens.		
Key		Location/Qualifiers		
PH		1..3489		
CDS				

FT		/tag= "a
XT		/product= "EMBR-1 protein"
XX		
PN	WO200248362-A2.	
PD		
XX	20-JUN-2002.	
PF	14-NOV-2001; 2001WO-US043956.	
XX		
PR	15-NOV-2000; 2000US-0249407P.	
XX	(INCYTE GENOMICS INC.	
PI	Ramkumar J, Arvizu C;	
XX		
DR	WPI; 2002-537629/57.	
DR	P-PEDB; AAB25053.	
PT	New polypeptides of human embryogenesis associated proteins for screening	
PT	modulators useful for treating or preventing disorders e.g.	
XX	endometriosis, infertility, allergy, preeclampsia.	
PS	Claim 58; Page 95-96; 97pp; English.	
XX		
CC	The invention relates to human embryogenesis associated proteins (EMBR)	
CC	and nucleic acid molecules encoding such proteins. EMBRY sequences are	
CC	useful for screening modulators useful for treating or preventing	
CC	disorders associated with abnormal expression of EMBRY. The disorders	
CC	treated include reproductive disorders such as infertility,	
CC	endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory	
CC	disorder such as acquired immune deficiency syndrome (AIDS), allergies,	
CC	contact dermatitis; disorders of the placenta such as preeclampsia,	
CC	abruptio placentae etc. Sequences of the invention are also useful for	
CC	analysing a proteome of a tissue or a cell type. EMBRY proteins are	
CC	useful as immunogens for preparing antibodies. Polynucleotides of the	
CC	invention are useful for creating knock-in humanised animals or transgenic	
CC	animals to model human diseases. They are also used in gene therapy. The	
CC	present sequence is human EMBRY-1 cDNA	
SQ	Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;	
Query Match	72.3%; Score 2917; DB 6; Length 3489;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 3017; Conservative	0; Mismatches 2; Indels 0; Gaps 0	
OY	290 AAATTTCACAAGCTATGGAAACAGAAGGTGCGCACGACGACGACAGACAGAAACAAGAA 349	
DB	452 AAAATTTCACAAGCTATGGAAACAGAAGGTGCGCACGACGACGACAGACAGAAACAAGAA 511	
OY	350 TTTCACAAGCTATGGAAACAGAAGGTGCGCACGACGACGACAGACAGAAACAAGGCATG 409	
DB	512 TTTCACAAGCTATGGAAACAGAAGGTGCGCACGACGACGACAGACAGAAACAAGGCATG 571	
OY	410 GAGGTGACACATGGGACTTACCAAGATCAGTGATGCCAATAATTCGCTGAGAGAGAGATG 469	
DB	572 GAGGTGACACATGGGACTTACCAAGATCAGTGATGCCAATAATTCGCTGAGAGAGAGATG 631	
OY	470 TAGCTGCGTAGTTTGAANAACA CTGCTGTGACTGGCCGGAATGAAACGTTGGCTGGTG 529	
DB	632 TAGCTGCGTAGTTTGAANAACA CTGCTGTGACTGGCCGGAATGAAACGTTGGCTGGTG 691	
OY	530 CTTTGTGATTTCAGACCGGTGGGGCTTCCGGGCTTCGACCGGTGGTTTCTGCACGGAAGATCAG 589	
DB	692 CTTTGTGATTTCAGACCGGTGGGGCTTCCGGGCTTCGACCGGTGGTTTCTGCACGGAAGATCAG 751	
OY	590 GAATTGGGAAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGACCTCT 649	
DB	752 GAATTGGGAAATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGACCTCT 811	
OY	650 ACCAGGGAATGTCTCTTAAGCTCTTCTTCTCCCGTTAAGAGATGACGCGGAAGGAAGG 709	
DB	812 ACCAGGGAATGTCTCTTAAGCTCTTCTTCTCCCGTTAAGAGATGACGCGGAAGGAAGG 871	

QY 710 AGACAGTGTCA CAGAGTTCA TCTCCAGGAGATG GGCACAGCTCCAGGCTCCGGTACGG 769  
| | | | |  
Db 872 AAGAGAGTGTCA CAGAGTTCA TCTCCAGGAGATG GGCACAGCTCCAGGCTCCGGTACGG 931  
| | | | |  
QY 770 AATATATGTCTCCGAC CAGAAAGGCTGTTGTTCA TCAATTGACCGGTTTCAGTGA CTTGAGGCT 829  
| | | | |  
Db 932 AATATATGTCTCCGAC CAGAAAGGCTGTTGTTCA TCAATTGACCGGTTTCAGTGA CTTGAGGCT 991  
| | | | |  
QY 830 CTGTCTCTCAACAA TGAACAAGAGCTGTGAAAGAC TGGGGCTGAAAGACGCTCCGTTCA 889  
| | | | |  
Db 992 CTGTCTCTCAACAA TGAACAAGAGCTGTGAAAGAC TGGGGCTGAAAGACGCTCCGTTCA 1051  
| | | | |  
QY 890 CCTCTATACGAGTGT GCTGAGAGAGGCTGTGCTCCCTGAGTCTTCCGTTACCTGATGTCACCG 949  
| | | | |  
Db 1052 CCTCTATACGAGTGT GCTGAGAGAGGCTGTGCTCCCTGAGTCTTCCGTTACCTGATGTCACCG 1111  
| | | | |  
QY 950 TCAGAGACGTGGGCA CAGAGAGCTCAAGTCAAGAGTCTGTCTCTCCGTTACCTGTTAG 1009  
| | | | |  
Db 1112 TCAGAGACGTGGGCA CAGAGAGCTCAAGTCAAGAGTCTGTCTCTCCGTTACCTGTTAG 1171  
| | | | |  
QY 1010 TTAGAGGAATCTCT GCGGGGAA CAAAGATCA CTTGCTCTTTAGCGCGGGATTTGGTACG 1069  
| | | | |  
Db 1172 TTAGAGGAATCTCT GCGGGGAA CAAAGATCA CTTGCTCTTTAGCGCGGGATTTGGTACG 1231  
| | | | |  
QY 1070 ATCAGAGACAGAA GGGTTGCGGCGATCATGAA CAAACGTTAGAGTGTCTGACCAAGTGC 1129  
| | | | |  
Db 1232 ATCAGAGACAGAA GGGTTGCGGCGATCATGAA CAAACGTTAGAGTGTCTGACCAAGTGC 1291  
| | | | |  
QY 1130 AAGTGCCTGCGCTG TGGGCTCTCTATCTGCGTGGCCCTGAGCTGACAGAGCTGTGAGG 1189  
| | | | |  
Db 1292 AAGTGCCTGCGCTG TGGGCTCTCTATCTGCGTGGCCCTGAGCTGACAGAGCTGTGAGG 1351  
| | | | |  
QY 1190 AAGAGGTGCGCCCT TTTCAACCAAGCTCA CAGGCTTGCACGCGCTTTTGGTTTCA TC 1249  
| | | | |  
Db 1352 AAGAGGTGCGCCCT TTTCAACCAAGCTCA CAGGCTTGCACGCGCTTTTGGTTTCA TC 1411  
| | | | |  
QY 1250 AACTCAACCCCTGAG GGGCGTGTCCGGCGCTGCTCAATCTGAGGAGAAAGATTGTCCTGA 1309  
| | | | |  
Db 1412 AACTCAACCCCTGAG GGGCGTGTCCGGCGCTGCTCAATCTGAGGAGAAAGATTGTCCTGA 1471  
| | | | |  
QY 1310 AAGCGCTTTCGCGAT GATGCTGTGAGAGGAGTGTGAATAGGAAGTCAAGTTTGAATGCTG 1369  
| | | | |  
Db 1472 AAGCGCTTTCGCGAT GATGCTGTGAGAGGAGTGTGAATAGGAAGTCAAGTTTGAATGCTG 1531  
| | | | |  
QY 1370 ACGAAGCTATGATG TCAAGGACTCTGGGAGATCTGAGAGTCTGCTGTCTGTTTCA CATTGA CA 1429  
| | | | |  
Db 1532 ACGAAGCTATGATG TCAAGGACTCTGGGAGATCTGAGAGTCTGCTGTCTGTTTCA CATTGA CA 1591  
| | | | |  
QY 1430 TCCCTCTCCGAGAC AGCCCATCTGTGAGAGATCTA CACCTTTTCCACCTCAAGTCTTCAGG 1489  
| | | | |  
Db 1592 TCCCTCTCCGAGAC AGCCCATCTGTGAGAGATCTA CACCTTTTCCACCTCAAGTCTTCAGG 1651  
| | | | |  
QY 1490 AACTCTGATGCGGCT TTTGACTACGTTGTAGAGGAGCTGTGAAATTCAGAGCAGGCTCTCTGCG 1549  
| | | | |  
Db 1652 AACTCTGATGCGGCT TTTGACTACGTTGTAGAGGAGCTGTGAAATTCAGAGCAGGCTCTCTGCG 1711  
| | | | |  
QY 1550 CTCTGTACGTTGA GAACAAGAGGCTCATGAGAGCTTAA CAGAGGCTTCATATCC 1609  
| | | | |  
Db 1712 CTCTGTACGTTGA GAACAAGAGGCTCATGAGAGCTTAA CAGAGGCTTCATATCC 1771  
| | | | |  
QY 1610 ACTGCGTTTGAAGA GCGTTTCTTGTGCTGTGAGCGAAGACGTTAAAGAGGCGCAC 1669  
| | | | |  
Db 1772 ACTGCGTTTGAAGA GCGTTTCTTGTGCTGTGAGCGAAGACGTTAAAGAGGCGCAC 1831  
| | | | |  
QY 1670 TGGAGGCTCTGCTG GAGCTGTCCCGTCCCTGAGGAGTGAAGAGCTTTCGACCTGAG 1729  
| | | | |  
Db 1832 TGGAGGCTCTGCTG GAGCTGTCCCGTCCCTGAGGAGTGAAGAGCTTTCGACCTGAG 1891  
| | | | |  
QY 1730 TCTCTCTGTTGGGT GATGAGAGCTTAATGCA CCACTCCAGAGACACCTCTGACGCTTCC 1789  
| | | | |  
Db 1892 TCTCTCTGTTGGGT GATGAGAGCTTAATGCA CCACTCCAGAGACACCTCTGACGCTTCC 1951  
| | | | |  
QY 1790 ACTGCTTTTGAAG ACTCAAGACAAAGATTGTTGCTGCTTGGCATTTAAACAGCTTCCAG 1849  
| | | | |

Db 1952 ACTGCTTTTGAAG ACTCAAGCAAAAGATTGTTGCTGTGGCATTTAAACAGCTTCCAG 2011  
| | | | |  
QY 1850 AAGTGTGCTTCGAT TAACCAAGACCTGAGCTTAATGACATCTTCTCTGCTCCAGC 1909  
| | | | |  
Db 2012 AAGTGTGCTTCGAT TAACCAAGACCTGAGCTTAATGACATCTTCTCTGCTCCAGC 2071  
| | | | |  
QY 1910 ACTGTCCGTAATTTG GGGGAAAAATTCGGGTGAGATGTCAAGAGGATCTTCCCAAGATGAGT 1969  
| | | | |  
Db 2072 ACTGTCCGTAATTTG GGGGAAAAATTCGGGTGAGATGTCAAGAGGATCTTCCCAAGATGAGT 2131  
| | | | |  
QY 1970 CCGTGAAGGATGTCT GGTGCTCCCTATATGAGATCGGGATTAAGACCTCATTTGAGAGC 2029  
| | | | |  
Db 2132 CCGTGAAGGATGTCT GGTGCTCCCTATATGAGATCGGGATTAAGACCTCATTTGAGAGC 2191  
| | | | |  
QY 2030 AAGTGAAGAAATTT GCTGCTTCATGCTTGGACCCACCA CACCTGTGGGAGCTGACCTGG 2089  
| | | | |  
Db 2192 AAGTGAAGAAATTT GCTGCTTCATGCTTGGACCCACCA CACCTGTGGGAGCTGACCTGG 2251  
| | | | |  
QY 2090 GAGCAGCATCTGTA CAGAGCGGGGCTATGAAAGACCTGTGTGTCGAAGCTGAGGCTCCA 2149  
| | | | |  
Db 2252 GAGCAGCATCTGTA CAGAGCGGGGCTATGAAAGACCTGTGTGTCGAAGCTGAGGCTCCA 2311  
| | | | |  
QY 2150 CCTGCAAGATACAG ACCCTGATGTTAGAAATGCA CAGATTA CCGCTGGTGTGACAGACC 2209  
| | | | |  
Db 2312 CCTGCAAGATACAG ACCCTGATGTTAGAAATGCA CAGATTA CCGCTGGTGTGACAGACC 2371  
| | | | |  
QY 2210 TCTGAGAAATCGTCA TGGGCCAACCGTTAACCTTAAGATCCCTCAACTTGGAGGACACCA CC 2269  
| | | | |  
Db 2372 TCTGAGAAATCGTCA TGGGCCAACCGTTAACCTTAAGATCCCTCAACTTGGAGGACACCA CC 2431  
| | | | |  
QY 2270 TGAAGAAAGAGATG TTAAGATGAGGTGAGTGAAGCTTTAAACACCAAAATGTTGTTGG 2329  
| | | | |  
Db 2432 TGAAGAAAGAGATG TTAAGATGAGGTGAGTGAAGCTTTAAACACCAAAATGTTGTTGG 2491  
| | | | |  
QY 2330 AGTCTTTGAGGCTGAT GTGCTGTGATGACCCATAGCCTGTTACCTGAGATCTCCCAA 2389  
| | | | |  
Db 2492 AGTCTTTGAGGCTGAT GTGCTGTGATGACCCATAGCCTGTTACCTGAGATCTCCCAA 2551  
| | | | |  
QY 2390 TCCCTTAAGACCTC CCCCAGCCCTGAATCTCTGAGCCTGAGCAAGAAACAAGTGA CAGACC 2449  
| | | | |  
Db 2552 TCCCTTAAGACCTC CCCCAGCCCTGAATCTCTGAGCCTGAGCAAGAAACAAGTGA CAGACC 2611  
| | | | |  
QY 2450 AAGGAGTAATGCTCT CAGTGAATGCTTTGAGAGTCTCCCAAGTGCCTGCAAGAGCTGA 2509  
| | | | |  
Db 2612 AAGGAGTAATGCTCT CAGTGAATGCTTTGAGAGTCTCCCAAGTGCCTGCAAGAGCTGA 2671  
| | | | |  
QY 2510 TACTGAGAGACGTGT GGCATCAAGCAGAGGGGTGCAAGATCTGGCCTCAAGCCCTGCA 2569  
| | | | |  
Db 2672 TACTGAGAGACGTGT GGCATCAAGCAGAGGGGTGCAAGATCTGGCCTCAAGCCCTGCA 2731  
| | | | |  
QY 2570 GCAACCGAGGCTTGA CACACCTGTGCTATCCACCAAGAGCTGTGGGAAACGAAGGTGTA 2629  
| | | | |  
Db 2732 GCAACCGAGGCTTGA CACACCTGTGCTATCCACCAAGAGCTGTGGGAAACGAAGGTGTA 2791  
| | | | |  
QY 2630 ATCTACTGTGATCA TGAAGGCTTCCCACTGTAAGTGTGCAAGGCTGATGCTGAATC 2689  
| | | | |  
Db 2792 ATCTACTGTGATCA TGAAGGCTTCCCACTGTAAGTGTGCAAGGCTGATGCTGAATC 2851  
| | | | |  
QY 2690 AGTGCACCTGAGAC AGGCTGCTGCTGCTGCTTTTCTTGA CTTGCGCTTAAGGTTACTAT 2749  
| | | | |  
Db 2852 AGTGCACCTGAGAC AGGCTGCTGCTGCTGCTTTTCTTGA CTTGCGCTTAAGGTTACTAT 2911  
| | | | |  
QY 2750 GAGTGAAGCACTGAG CTTAGATGAACCCCTGTGGAAGACAAATGAGCTGAAGCTTCTGT 2809  
| | | | |  
Db 2912 GAGTGAAGCACTGAG CTTAGATGAACCCCTGTGGAAGACAAATGAGCTGAAGCTTCTGT 2971  
| | | | |  
QY 2810 GCGAGGTATGAGAA CCAATCTTGTATCTCAGAGACCTGGAAGTTGGTAAAGTCTATC 2869  
| | | | |  
Db 2972 GCGAGGTATGAGAA CCAATCTTGTATCTCAGAGACCTGGAAGTTGGTAAAGTCTATC 3031  
| | | | |  
QY 2870 TCACCGCGCGGTGTA GAGTCTGTCTGTGATCTTCAGAGAGGAGACACCTGAAGA 2929  
| | | | |



Db 1172 TTAGAGAAATCTCCGGGAAACAAAGAAATCCATTGCTCTTGAGCGGGGATTTGTGAGC 1231  
Qy 1070 ATCAGAAACACAAAGGTTGCGTGGATATATGAACAACCTGAGCTGCTCAACAGTGC 1129  
Db 1232 ATCAGAAACACAAAGGTTGCGTGGATATATGAACAACCTGAGCTGCTCAACAGTGC 1291  
Qy 1130 AGGTGCGCGCGTGGGCTCTCTCATCTGCGGCGCTGAGCGGAGAGAGTGTGGGG 1189  
Db 1292 AGGTGCGCGCGTGGGCTCTCTCATCTGCGGCGCTGAGCGGAGAGAGTGTGGGG 1351  
Qy 1190 AGAGCGCGCGCTCTCAACAACAACGCTCAAGCGCTGACCGCTTTTGTGTTCAATC 1249  
Db 1352 AGAGCGCGCGCTCTCAACAACAACGCTCAAGCGCTGACCGCTTTTGTGTTCAATC 1411  
Qy 1250 AGCTCACTCTTGAGGCGTGTGCGCGCTGTCTCATTTGAGAGAAAGTGTCTGA 1309  
Db 1412 AGCTCACTCTTGAGGCGTGTGCGCGCTGTCTCATTTGAGAGAAAGTGTCTGA 1471  
Qy 1310 AGCGCTTCTGCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAATGTTGATGTTG 1369  
Db 1472 AGCGCTTCTGCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAATGTTGATGTTG 1531  
Qy 1370 AGACCTCATGATTAAGAGACTCGGAGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1429  
Db 1532 AGACCTCATGATTAAGAGACTCGGAGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1591  
Qy 1430 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTCCACTCACTTCAGG 1489  
Db 1592 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTCCACTCACTTCAGG 1651  
Qy 1490 ACTCTGTGCGCGCTTGTATCTACGTGTGAGAGGCGTGAAGTCCAGCGAGCTCTGCG 1549  
Db 1652 ACTCTGTGCGCGCTTGTATCTACGTGTGAGAGGCGTGAAGTCCAGCGAGCTCTGCG 1711  
Qy 1550 CTCTGATCTTGAAGAACAAGAGTCCATGAGGCTTAACAGGAGGCTTCCATATCC 1609  
Db 1712 CTCTGATCTTGAAGAACAAGAGTCCATGAGGCTTAACAGGAGGCTTCCATATCC 1771  
Qy 1610 ACTCGCTTGTGATGAAGCGCTTTCTGTTTGGCCTGTGAGCGAAGCGTAAGAGGCGAC 1669  
Db 1772 ACTCGCTTGTGATGAAGCGCTTTCTGTTTGGCCTGTGAGCGAAGCGTAAGAGGCGAC 1831  
Qy 1670 TGGAGGTCCGCGGCTGTCCCGTTCCGCTGGGGGTGAAGAGAGCTTCTGCACTGGG 1729  
Db 1832 TGGAGGTCCGCGGCTGTCCCGTTCCGCTGGGGGTGAAGAGAGCTTCTGCACTGGG 1891  
Qy 1730 TCTCTCTGTGGGATCAGACCTTAATGCAACCAACCCAGAGACAACCTGAGCGCTTCC 1789  
Db 1892 TCTCTCTGTGGGATCAGACCTTAATGCAACCAACCCAGAGACAACCTGAGCGCTTCC 1951  
Qy 1790 ACTGTCTTGTGAGACTCAAGACAAGAGTTGTGCTTGGCATTAAACAGCTTCCAG 1849  
Db 1952 ACTGTCTTGTGAGACTCAAGACAAGAGTTGTGCTTGGCATTAAACAGCTTCCAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAACAGAACTTGAATACATCTTCTTCTGCTTCCAGC 2071  
Db 2012 AAGTGTGCTTCCGATTAACAGAACTTGAATACATCTTCTTCTGCTTCCAGC 2071  
Qy 1910 ACTGTCCGATTTTGGAGAAATTCGGGTGATGTCAAAAGAGATCTTCCAGAGATGAGT 1969  
Db 2072 ACTGTCCGATTTTGGAGAAATTCGGGTGATGTCAAAAGAGATCTTCCAGAGATGAGT 2131  
Qy 1970 CCGCTGAGGCAATGTCTGTGTCTCTTATGAGATGCGGATTAAGACCTCATTTAGAGAC 2029  
Db 2132 CCGCTGAGGCAATGTCTGTGTCTCTTATGAGATGCGGATTAAGACCTCATTTAGAGAC 2191  
Qy 2030 AAGTGGAAATTTTGTCTTCAATGCTTGGACCAACCAACCTGCGGAGCTGAGACTGG 2089  
Db 2192 AAGTGGAAATTTTGTCTTCAATGCTTGGACCAACCAACCTGCGGAGCTGAGACTGG 2251  
Qy 2090 GGAGCAGATCTGACAGAGCGGAGCATGAAGACCTGTGTGCAAGAGTGAAGATCCA 2149  
Db 2252 GGAGCAGATCTGACAGAGCGGAGCATGAAGACCTGTGTGCAAGAGTGAAGATCCA 2311

Qy 2150 CTTGCAAGATACAGACCTGTGATTTAGAAATGCAAGATTAACCTCTGTGTGAGACAC 2209  
Db 2312 CTTGCAAGATACAGACCTGTGATTTAGAAATGCAAGATTAACCTCTGTGTGAGACAC 2371  
Qy 2210 TCTGAGATAGTCAATGCGCAACCGTAACCTTAAGTCCCTGAACCTTGGAGGCAACCC 2269  
Db 2372 TCTGAGATAGTCAATGCGCAACCGTAACCTTAAGTCCCTGAACCTTGGAGGCAACCC 2431  
Qy 2270 TGAAGAAAGAGATTAAGATGCGTGTGAAGCTTAAACCAACCAAAATGTTGTTGG 2329  
Db 2432 TGAAGAAAGAGATTAAGATGCGTGTGAAGCTTAAACCAACCAAAATGTTGTTGG 2491  
Qy 2330 AGCTTTAAGCTGATTTGCTGTGATGACCAATGCTGTGTAACCTGAAGATCTCCAAA 2389  
Db 2492 AGCTTTAAGCTGATTTGCTGTGATGACCAATGCTGTGTAACCTGAAGATCTCCAAA 2551  
Qy 2390 TCTTACGACCTCCCGCAGCTGAATCTTGAGCCTGAGAGAAACAAAGTGAAGAC 2449  
Db 2552 TCTTACGACCTCCCGCAGCTGAATCTTGAGCCTGAGAGAAACAAAGTGAAGAC 2611  
Qy 2450 AAGGATATGCTCTCAATGATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2509  
Db 2612 AAGGATATGCTCTCAATGATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGAGATCAACAGCAGGCTGCAAGTCTGAGCCTGCTGCA 2569  
Db 2672 TACTGAGAGACTGTGAGATCAACAGCAGGCTGCAAGTCTGAGCCTGCTGCA 2731  
Qy 2570 GCAACCGGAGCTTGACACACCTGTGCTATCAACCAACAGCTGGGAAACGAAGGTGA 2629  
Db 2732 GCAACCGGAGCTTGACACACCTGTGCTATCAACCAACAGCTGGGAAACGAAGGTGA 2791  
Qy 2630 ATCTACTGTGATCATGAGGCTTCCCACTGATCTGCAAGGCTGATGCTGAATC 2689  
Db 2792 ATCTACTGTGATCATGAGGCTTCCCACTGATCTGCAAGGCTGATGCTGAATC 2851  
Qy 2690 AAGTCACTCTGACACCGCTGCTGTGTTTCTTGGCACTTGCCTTATGGGATCAT 2749  
Db 2852 AAGTCACTCTGACACCGCTGCTGTGTTTCTTGGCACTTGCCTTATGGGATCAT 2911  
Qy 2750 GAGTCAACGCACTGAGCTTATGATGAACCTGTGGAAGACAATGCGCTGAAGCTTCTGT 2809  
Db 2912 GAGTCAACGCACTGAGCTTATGATGAACCTGTGGAAGACAATGCGCTGAAGCTTCTGT 2971  
Qy 2810 GCGAGTATGAGAGAACATCTTGTATCTTCAAGACCTGGAAGTTGTAAGTGCATC 2869  
Db 2972 GCGAGTATGAGAGAACATCTTGTATCTTCAAGACCTGGAAGTTGTAAGTGCATC 3031  
Qy 2870 TCAACCGCGCGTGTGTGAGAGTCTGCTGTGTGATCTGAGAGCAACCTGAAGA 2929  
Db 3032 TCAACCGCGCGTGTGTGAGAGTCTGCTGTGTGATCTGAGAGCAACCTGAAGA 3091  
Qy 2930 GCTGTGATCTACGAGCAATGCGCTGAGTGAACGCTGAGGCTTGTGCGAGGAGC 2989  
Db 3092 GCTGTGATCTACGAGCAATGCGCTGAGTGAACGCTGAGGCTTGTGCGAGGAGC 3151  
Qy 2990 TGAACCAAAAGAACAGTGTCTGACAGACTCGGTTGAAGGCAATGGAATGACTTCTG 3049  
Db 3152 TGAACCAAAAGAACAGTGTCTGACAGACTCGGTTGAAGGCAATGGAATGACTTCTG 3211  
Qy 3050 ATTGCTGTGAGGCACTTCTCTTGGCCCTTCTGCAACCGGATGACAGTCTAAAC 3109  
Db 3212 ATTGCTGTGAGGCACTTCTCTTGGCCCTTCTGCAACCGGATGACAGTCTAAAC 3271  
Qy 3110 TGTGCAAGATTAACCTTCAAGTCCAAAGGAATGATGAAGCTGTTGCGCTTGTGCTGTC 3169  
Db 3272 TGTGCAAGATTAACCTTCAAGTCCAAAGGAATGATGAAGCTGTTGCGCTTGTGCTGTC 3331  
Qy 3170 CCAAGCTTAACCTTCAAGTAAATGGGCTGTGGAATGCAAGTACCTGTGCAAAATGAAGA 3229  
Db 3332 CCAAGCTTAACCTTCAAGTAAATGGGCTGTGGAATGCAAGTACCTGTGCAAAATGAAGA 3391







XX OS unidentified.  
XX PN WO200240668-A2.  
XX PD 23-MAY-2002.  
XX PF 30-OCT-2001; 2001WO-BP012545.  
XX PR 15-NOV-2000; 2000DE-01056687.  
XX PR 30-NOV-2000; 2000DE-01059595.  
XX PA (ABOT-) APOTEC RES & DEV LTD.  
XX PI Tschoopp J, Martinon F;  
XX DR WPI; 2002-427093/45.  
XX DR P-PSDB; AA017863.  
XX PT New DNA encoding protein with pyrin domain, useful for treating diseases  
XX PT involving impaired signal transduction, particularly inflammation, also  
XX PT proteins and antibodies.  
XX PS Claim 5; Fig 1; 116pp; German.  
XX CC The present invention relates the DNA and their encoded proteins, where  
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used  
XX CC to treat diseases associated with impaired intracellular signal  
XX CC transduction, particularly inflammation such as psoriasis,  
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis  
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,  
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's  
XX CC and Parkinson's diseases. The present sequence is a coding sequence of  
XX CC the invention  
XX S0 Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;  
Query Match 69.5%; Score 2803; DB 6; Length 5859;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 290 AAATTTCACAACTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGAA 349  
DB 2387 AAATTTCACAACTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGAA 2446  
QY 350 TTTTCAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGAACTG 409  
DB 2447 TTTTCAAGCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAACAAGAACTG 2506  
QY 410 GAGGTGACATGGAACAAGAGGTGCTGATGACCAAAATGCGCTGAGAGAGAGATG 469  
DB 2507 GAGGTGACATGGAACAAGAGGTGCTGATGACCAAAATGCGCTGAGAGAGAGATG 2566  
QY 470 TACGTGATGTTTGAACAACTGCTGCTGACTGCGCGGAAATGCAACGTTGCTGCTG 529  
DB 2567 TACGTGATGTTTGAACAACTGCTGCTGACTGCGCGGAAATGCAACGTTGCTGCTG 2626  
QY 530 CTTTGTGATTCAGACCGGTGGGGCTTCCGGCTCGACCGGTGTTTGCACCGAAAGTCA 589  
DB 2627 CTTTGTGATTCAGACCGGTGGGGCTTCCGGCTCGACCGGTGTTTGCACCGAAAGTCA 2686  
QY 590 GAATTGGGAAATGCGCTTACGACAGAGATGCTGCTGCTGGCGGCAAGTGGACTCT 649  
DB 2687 GAATTGGGAAATGCGCTTACGACAGAGATGCTGCTGCTGGCGGCAAGTGGACTCT 2746  
QY 650 ACCAGGAGATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGCAAGCGAAAGAG 709  
DB 2747 ACCAGGAGATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGCAAGCGAAAGAG 2806  
QY 710 AGAGCAGTGTCAACAGAGTTTCACTTCAGAGAGTGGCCAGACTCCAGGCTCCGGTGAAG 769  
DB 2807 AGAGCAGTGTCAACAGAGTTTCACTTCAGAGAGTGGCCAGACTCCAGGCTCCGGTGAAG 2866

QY 770 AGATCATGTCCCGACAGAAAGCGTGTGTTTCATCATTTGACGGTTCGATGACTGGGCT 829  
DB 2867 AGATCATGTCCCGACAGAAAGCGTGTGTTTCATCATTTGACGGTTCGATGACTGGGCT 2926  
QY 830 CTGTCTCAGACATGACCAAAAGCTCTGCAAAAGCTGGGGCTGAGAGACAGCTCCGTTCA 889  
DB 2927 CTGTCTCAGACATGACCAAAAGCTCTGCAAAAGCTGGGGCTGAGAGACAGCTCCGTTCA 2986  
QY 890 CCTTCATGCGAGCTGTGAGGAAGGTCCTGCTCCGTAGTCTCTCTGATGCTGACCG 949  
DB 2987 CCTTCATGCGAGCTGTGAGGAAGGTCCTGCTCCGTAGTCTCTCTGATGCTGACCG 3046  
QY 950 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGGTGCTCTCCCGTTACCTGTTAG 1009  
DB 3047 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGGTGCTCTCCCGTTACCTGTTAG 3106  
QY 1010 TTAGAGGATTCCTGGGGACAAAGATCCACTTGTCTCTTGAAGCGGGATTTGGTAGC 1069  
DB 3107 TTAGAGGATTCCTGGGGACAAAGATCCACTTGTCTCTTGAAGCGGGATTTGGTAGC 3166  
QY 1070 ATCAGAGACACAGAGGTTGGTGGATGACCAACCGTGAAGCTGTCGACAGTGGC 1129  
DB 3167 ATCAGAGACACAGAGGTTGGTGGATGACCAACCGTGAAGCTGTCGACAGTGGC 3226  
QY 1130 AGGTGCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGACGTGGTGGGG 1189  
DB 3227 AGGTGCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGACGTGGTGGGG 3286  
QY 1190 AGAGGTGGCCCTCTTCAACCAACGCTCACAGGCTGACAGCGCTTTTGTGTTTATC 1249  
DB 3287 AGAGGTGGCCCTCTTCAACCAACGCTCACAGGCTGACAGCGCTTTTGTGTTTATC 3346  
QY 1250 AGCTCACCCCTGAGGGGTGGTCCGGCGTGTCTCAATCTGAGAGAAAGATTGCTGTA 1309  
DB 3347 AGCTCACCCCTGAGGGGTGGTCCGGCGTGTCTCAATCTGAGAGAAAGATTGCTGTA 3406  
QY 1310 AGCCCTTCTGCGCTATGAGCTGTGAGGAGAGTGGAAATGAGAACTAGTGTATGATG 1369  
DB 3407 AGCCCTTCTGCGCTATGAGCTGTGAGGAGAGTGGAAATGAGAACTAGTGTATGATG 3466  
QY 1370 ACGACCTCATGTTCAAGAGACTCGGGGAGTGTAGCTCCGTCTCTGTTTCACTGAACA 1429  
DB 3467 ACGACCTCATGTTCAAGAGACTCGGGGAGTGTAGCTCCGTCTCTGTTTCACTGAACA 3526  
QY 1430 TCCCTTCCGACAGACGACCTGTGAGAGTACTACACTTCTTCACTGACTCTCCAGG 1489  
DB 3527 TCCCTTCCGACAGACGACCTGTGAGAGTACTACACTTCTTCACTGACTCTCCAGG 3586  
QY 1490 ACTTCTGTGCCCTTGTACTACGTGTTAAGAGGCTGGAATGAGCCAGCTCTGCC 1549  
DB 3587 ACTTCTGTGCCCTTGTACTACGTGTTAAGAGGCTGGAATGAGCCAGCTCTGCC 3646  
QY 1550 CTCTGTACGTTGAGAGACAAAGAGTTCATGAGCTTTAAACAGGCAAGCTTCATATCC 1609  
DB 3647 CTCTGTACGTTGAGAGACAAAGAGTTCATGAGCTTTAAACAGGCAAGCTTCATATCC 3706  
QY 1610 ACTGCTTTGGATGAAGGTTCTTGTGTCCTGTGAGGAGGAGAGATGAAGAGAGGACAC 1669  
DB 3707 ACTGCTTTGGATGAAGGTTCTTGTGTCCTGTGAGGAGGAGAGATGAAGAGAGGACAC 3766  
QY 1670 TGAAGTCTGCTGGGCTGTCCGTTCCCTGAGGGGTGAAGCAGAACTTCTGACTGGG 1729  
DB 3767 TGAAGTCTGCTGGGCTGTCCGTTCCCTGAGGGGTGAAGCAGAACTTCTGACTGGG 3826  
QY 1730 TCTCTGTGTGGTCAAGACCTTAATGCAACACCCAGAGACACCTTGAGCCCTTCC 1789  
DB 3827 TCTCTGTGTGGTCAAGACCTTAATGCAACACCCAGAGACACCTTGAGCCCTTCC 3886  
QY 1790 ACTGCTTTTGAACCTCAAGACAAAGTGTGCTGCTGCACTTAAACAGCTTCCAG 1849  
DB 3887 ACTGCTTTTGAACCTCAAGACAAAGTGTGCTGCTGCACTTAAACAGCTTCCAG 3946  
QY 1850 AAGTGTGGCTTCCGATTAACAGAACTGGAATTGATGATCTTCTGCTCCAGC 1909



```

Db 3947 AAGTGGGCTTCGATTAACGAACTGGAATTGATGCACTTCCTTCTGCTCCAGC 4006
Qy 1910 ACTGTCCTATTTCGGGAAAATTCGGGTGATGTCMAAGGATCTTCCACAGATGAGT 1969
Db 4007 ACTGTCCTATTTCGGGAAAATTCGGGTGATGTCMAAGGATCTTCCACAGATGAGT 4066
Qy 1970 CCGGAGGACATGTCCTGTCCTCTATGGAATCGGGATGAAGCCCTCATTTGAGAGC 2029
Db 4067 CCGGAGGACATGTCCTGTCCTCTATGGAATCGGGATGAAGCCCTCATTTGAGAGC 4126
Qy 2030 AGTGGGAGATTTCTGCTCCATGCTTGGAACCCACCACTGCGGAGCTGACCTGG 2089
Db 4127 AGTGGGAGATTTCTGCTCCATGCTTGGAACCCACCACTGCGGAGCTGACCTGG 4186
Qy 2090 GCAGCAGCATCTCTGACAGAGCGGCGCATGAGAGCCCTGTGTCCAACTGAGGATCCCA 2149
Db 4187 GCAGCAGCATCTCTGACAGAGCGGCGCATGAGAGCCCTGTGTCCAACTGAGGATCCCA 4246
Qy 2150 CCTGGAATACAGACCTGATGTTTAAAGACAGATTACCCCTGTCGACGAC 2209
Db 4247 CCTGGAATACAGACCTGATGTTTAAAGACAGATTACCCCTGTCGACGAC 4306
Qy 2210 TCTGAGATCGTCATGCGCAACCGTAACTTAAGTCCCTCACTTGGGAGGACCCAC 2269
Db 4307 TCTGAGATCGTCATGCGCAACCGTAACTTAAGTCCCTCACTTGGGAGGACCCAC 4366
Qy 2270 TGAAGGAAAGATGTAAAGATGCGGTGAAGCCTTAAACACCCAAATGTTTGG 2329
Db 4367 TGAAGGAAAGATGTAAAGATGCGGTGAAGCCTTAAACACCCAAATGTTTGG 4426
Qy 2330 AGCTTTTGAAGGCTGATGCTGTGATGACCCATGCTGTTAAGTGAAGATCTCCCAA 2389
Db 4427 AGCTTTTGAAGGCTGATGCTGTGATGACCCATGCTGTTAAGTGAAGATCTCCCAA 4486
Qy 2390 TCTTAAAGACCTCCCCAGCCTGAATCTGTGAGCTGGCAGAGAAACAAGTGAAGACC 2449
Db 4487 TCTTAAAGACCTCCCCAGCCTGAATCTGTGAGCTGGCAGAGAAACAAGTGAAGACC 4546
Qy 2450 AAGGAGTAATGCTCTCAGTGAATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2509
Db 4547 AAGGAGTAATGCTCTCAGTGAATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 4606
Qy 2510 TACTGAGAGACTGTGAGCATCAAGCCAGGGTTGCCAAGATGTGGCCTTCAAGCTTCA 2569
Db 4607 TACTGAGAGACTGTGAGCATCAAGCCAGGGTTGCCAAGATGTGGCCTTCAAGCTTCA 4666
Qy 2570 GCAACCGGAGCTTGACACACCTGTGCTATCCAAACAAGCCTGGGGAACGAAGGTGTA 2629
Db 4667 GCAACCGGAGCTTGACACACCTGTGCTATCCAAACAAGCCTGGGGAACGAAGGTGTA 4726
Qy 2630 ATCTACTGTGTGATCATGAGGCTTCCCACTGATGCTGCAAGAGGCTGATGCTGAATC 2689
Db 4727 ATCTACTGTGTGATCATGAGGCTTCCCACTGATGCTGCAAGAGGCTGATGCTGAATC 4786
Qy 2690 AGTGCACCTTGACACCGGCTGCTGTGTTTCTTGCACTTGCGCTTATGGGTAATCAT 2749
Db 4787 AGTGCACCTTGACACCGGCTGCTGTGTTTCTTGCACTTGCGCTTATGGGTAATCAT 4846
Qy 2750 GGCCTGAGCACCCTGAGCCTTATGACGAACCCGTGGAAGACAAATGCGTGAAGCTTCTGT 2809
Db 4847 GGCCTGAGCACCCTGAGCCTTATGACGAACCCGTGGAAGACAAATGCGTGAAGCTTCTGT 4906
Qy 2810 GCGAGGTATGAGAGAACCATTTGTCTCATCTTCCAGACCTGAGATTGTAAGTGTCAATC 2869
Db 4907 GCGAGGTATGAGAGAACCATTTGTCTCATCTTCCAGACCTGAGATTGTAAGTGTCAATC 4966
Qy 2870 TCACCGCGCGTGTGCTGAGAGTCTGTCTGTGTGATCTCGAGAGCAGACACTGAAGA 2929
Db 4967 TCACCGCGCGTGTGCTGAGAGTCTGTCTGTGTATCTCGAGAGCAGACACTGAAGA 5026
Qy 2930 GCGTGAATCTCAACGACAAATGCGGTGGAAGCGGTGGTTCGCTGCGGAGGAC 2989

```

```

Db 5027 GCTGTGATCTCAGGACAAATGCCCTGGGTGAACGGTGGGTTGTGCACTGTGCGAGGAC 5086
Qy 2990 TGAACCAAAAGAACAGTGTCTTGACAGACTCGGGTTGAAGGCAATGAGACTGTCTG 3049
Db 5087 TGAACCAAAAGAACAGTGTCTTGACAGACTCGGGTTGAAGGCAATGAGACTGTCTG 5146
Qy 3050 ATTGCTGTGAGGACATCTCTCTGAGCCCTTCTGCAACCGGCACTGACCAAGTCTAAAC 3109
Db 5147 ATTGCTGTGAGGACATCTCTCTGAGCCCTTCTGCAACCGGCACTGACCAAGTCTAAAC 5206
Qy 3110 TGTGTCAAGATTAATCTTCAAGTCCCAAGAGATGATGAAGCTGTTCGGCTTGTCTGTC 3169
Db 5207 TGTGTCAAGATTAATCTTCAAGTCCCAAGAGATGATGAAGCTGTTCGGCTTGTCTGTC 5266
Qy 3170 CCAGGTCTTAATCTTACAGTAATTTGG 3194
Db 5267 CCAGGTCTTAATCTTACAGTAATTTGG 5291

```

```

RESULT 7
AAI47131
ID AAI47131 standard; DNA; 6939 BP.
XX
XX AAI47131;
XX
XX 20-AUG-2002 (first entry)
XX
XX Pyrin domain containing protein NALP5/Py8-hs coding sequence.
DE
XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
XX antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
XX neuroprotective; antiarthritic; antirheumatic; antineumatic;
XX nephroprotective; osteoprotective; nootropic; intracellular signal transduction;
XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;
XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX osteoarthritis; glomerulonephritis; gene; ds.
XX
XX Unidentified.
OS
XX
XX W0200240668-A2.
XX
XX 23-MAY-2002.
XX
XX 30-OCT-2001; 2001WO-EP012545.
XX
XX 15-NOV-2000; 2000DE-01056687.
XX
XX 30-NOV-2000; 2000DE-01059595.
XX
XX (APOT-) APOTTECH RES & DEV LTD.
XX
XX Tschoep J, Martinon F;
XX
XX WPI; 2002-427093/45.
XX
XX P-PsDB; AAO17859.
XX
XX New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX
XX Claim 5; Fig 1; 116pp; German.
XX
XX The present invention relates the DNA and their encoded proteins, where
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used
XX CC to treat diseases associated with impaired intracellular signal
XX CC transduction, particularly inflammation such as psoriasis,
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX CC and Parkinson's diseases. The present sequence is a coding sequence of
XX CC the invention.
XX
XX Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX

```



```

Db 5567 TCCTTACGACCTCCCGCAGCCTGAAATCTCTGAGCCTGGCAGAGAAACAAGTGACAGACC 5626
Qy 2450 AGGAGTAATGCTCTGAGTATGCTTGAAGTCTCCAGTGGCGCCCTGAGAAAGCTGA 2509
Db 5627 AGGAGTAATGCTCTGAGTATGCTTGAAGTCTCCAGTGGCGCCCTGAGAAAGCTGA 5686
Qy 2510 TACTGAGAGCTGTGGCATCAGGCCACGGGTGGCCAGATCTGGCTTCAGCCCTGTCA 2569
Db 5687 TACTGAGAGCTGTGGCATCAGGCCACGGGTGGCCAGATCTGGCTTCAGCCCTGTCA 5746
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTATCCAAACAAGCCTGGGAAACGAAGGTGTA 2629
Db 5747 GCAACCGGAGCTTGAACAACCTGTGCTATCCAAACAAGCCTGGGAAACGAAGGTGTA 5806
Qy 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGATGATGTGAGAGGCTGATGCTGAATC 2689
Db 5807 ATCTACTGTGTGATCCATGAGGCTTCCCACTGATGATGTGAGAGGCTGATGCTGAATC 5866
Qy 2690 AGTGCACCTGAGACAGCGCTGCTGTGTTTCTTGACATTCGCTTATGGGTAACTCAT 2749
Db 5867 AGTGCACCTGAGACAGCGCTGCTGTGTTTCTTGACATTCGCTTATGGGTAACTCAT 5926
Qy 2750 GAGTCACGACCTGAGCCTTACATGAACCTGTGGAACAATGGCGTGAAGCTTCTGT 2809
Db 5927 GAGTCACGACCTGAGCCTTACATGAACCTGTGGAACAATGGCGTGAAGCTTCTGT 5986
Qy 2810 GCGAGCTCATGAGAAACATTTGTTCATCTCCAGAGCCTGAGATTTGGTAAAGTCTATC 2869
Db 5987 GCGAGCTCATGAGAAACATTTGTTCATCTCCAGAGCCTGAGATTTGGTAAAGTCTATC 6046
Qy 2870 TCACGCGCGCTGTGTGAGATCTGTCTGTGTATCTCGAGAGACAGACCTGGAAGA 2929
Db 6047 TCACGCGCGCTGTGTGAGATCTGTCTGTGTATCTCGAGAGACAGACCTGGAAGA 6106
Qy 2930 GCTTGAATCTCAGCAATGCTGCTGTGAGCAGTGGGTGCTGCTGTGCGAGGAGC 2989
Db 6107 GCTTGAATCTCAGCAATGCTGCTGTGAGCAGTGGGTGCTGCTGTGCGAGGAGC 6166
Qy 2990 TGAAGCAAAAGAAAGTGTTCAGAGACCTGGGTGAAGGCAATGTGACCTTCTG 3049
Db 6167 TGAAGCAAAAGAAAGTGTTCAGAGACCTGGGTGAAGGCAATGTGACCTTCTG 6226
Qy 3050 ATTTGCTGAGGACCTCTCTTGGGCTTCTGCAACCGGATCGACATCTAAACC 3109
Db 6227 ATTTGCTGAGGACCTCTCTTGGGCTTCTGCAACCGGATCGACATCTAAACC 6286
Qy 3110 TGGTCAGAAATTAATCTTCAAGTCCCAAGATGATGAGCTGTGCTGCTTGGCTGTC 3169
Db 6287 TGGTCAGAAATTAATCTTCAAGTCCCAAGATGATGAGCTGTGCTGCTTGGCTGTC 6346
Qy 3170 CCACGCTTAACTTACAGATTAATTGG 3194
Db 6347 CCACGCTTAACTTACAGATTAATTGG 6371

```

RESULT 8  
 AAL47140  
 ID AAL47140 standard; DNA; 6939 BP.

XX AAL47140;  
 XX 20-AUG-2002 (first entry)  
 DE Pyrin domain containing protein NALP3/Py17 coding sequence.  
 XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;  
 XX antiarteriosclerotic; antiposoriatic; antibacterial; virucide;  
 XX neuroprotective; antiarthritic; antirheumatic; antiaschematic;  
 XX nephroprotective; osteoprotective; neurotropic; intracellular signal transduction;  
 XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
 XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;  
 XX osteoarthritis; glomerulonephritis; gene; ds.

```

OS Unidentified.
XX
XX MO200240668-A2.
XX
XX 23-MAY-2002.
XX
XX 30-OCT-2001; 2001MO-EP012545.
XX
XX 15-NOV-2000; 2000DE-01056687.
XX
XX 30-NOV-2000; 2000DE-01059595.
XX
XX (APOT-) APOTEC RES & DEV LTD.
XX
XX Tschoep J, Martimon F;
XX
XX WPI: 2002-427093/45.
XX
XX P-PSDB; AA017868.
XX
XX
XX New DNA encoding protein with pyrin domain, useful for treating diseases
XX involving impaired signal transduction, particularly inflammation, also
XX proteins and antibodies.
XX
XX Claim 5; Fig 1; 116pp; German.
XX
XX The present invention relates the DNA and their encoded proteins, where
XX the proteins contain at least one PYD (pyrin) domain. These can be used
XX to treat diseases associated with impaired intracellular signal
XX transduction, particularly inflammation such as psoriasis,
XX arteriosclerosis, bacterial or viral infections (particularly meningitis
XX and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX the invention
XX
XX Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX Query Match 69.5%; Score 2803; DB 6; Length 6939;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 2003; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
Qy 290 AAATTTCAAGCTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAGAA 349
Db 3467 AAATTTCAAGCTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAGAA 3526
Qy 350 TTTCAAGCTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAGAA 409
Db 3527 TTTCAAGCTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAGAA 3586
Qy 410 GAGTGACACATGGGACCTACAGAGTCACTGATGACCAAAATTCGCTGAGAGAGATG 469
Db 3587 GAGTGACACATGGGACCTACAGAGTCACTGATGACCAAAATTCGCTGAGAGAGATG 3646
Qy 470 TACGCTAGATTTTGAAGAAACCTGCTGAGCTGGCCGGAATGCAAGCTTGGCTG 529
Db 3647 TACGCTAGATTTTGAAGAAACCTGCTGAGCTGGCCGGAATGCAAGCTTGGCTG 3706
Qy 530 CTTTGTATCAGACCGGTGGGGCTTCCGACCTGCGACGCTGCTTCTGACGGAAGTCA 589
Db 3707 CTTTGTATCAGACCGGTGGGGCTTCCGACCTGCGACGCTGCTTCTGACGGAAGTCA 3766
Qy 590 GAATTTGGGAATTCGCTTACGCAAGAGATCTGTGTGTGGCGGCAAGTGAATCT 649
Db 3767 GAATTTGGGAATTCGCTTACGCAAGAGATCTGTGTGTGGCGGCAAGTGAATCT 3826
Qy 650 ACCAGGAATGTTCTCTACGCTTCTTCTCCCGTTAGAGAGATGACGGAAGAAAG 709
Db 3827 ACCAGGAATGTTCTCTACGCTTCTTCTCCCGTTAGAGAGATGACGGAAGAAAG 3886
Qy 710 AGAGAGTGTCAAGAGTTCTTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTGACGG 769
Db 3887 AGAGAGTGTCAAGAGTTCTTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTGACGG 3946
Qy 770 AGATCATGTCCCGACAGAAAGGCTGTGTTATCATTTGACGCTTTCGATACCTGGGCT 829

```

Db 3947 AGATCATATGCCGACCAAGAAAGGCTGTTTCATTCATGACGGTTTCATGACCTCGGCT 4006  
Qy 830 CTGTCTTAACAAGACAAAGCTCTGCAGAAAGATCGGGCTGAGAGACCTCCGTTCA 889  
Db 4007 CTGTCTTAACAAGACAAAGCTCTGCAGAAAGATCGGGCTGAGAGACCTCCGTTCA 4066  
Qy 890 CCTCATACGAGAGTGTGCTGAGAGAGTCTGTCTCCCTGAGTCTTCTGTATCTGACCG 949  
Db 4067 CCTCATACGAGAGTGTGCTGAGAGAGTCTGTCTCCCTGAGTCTTCTGTATCTGACCG 4126  
Qy 950 TCAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAG 1009  
Db 4127 TCAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAG 4186  
Qy 1010 TTAGAGAAATCTCCGGGAAACAAGATTCATCTGCTCTTGAAGCGGGATTGTGAGC 1069  
Db 4187 TTAGAGAAATCTCCGGGAAACAAGATTCATCTGCTCTTGAAGCGGGATTGTGAGC 4246  
Qy 1070 ATCAGAAACAAGAGGTTGCTGATGAAACAACGTTGAGCTGCTGACCAAGTGC 1129  
Db 4247 ATCAGAAACAAGAGGTTGCTGATGAAACAACGTTGAGCTGCTGACCAAGTGC 4306  
Qy 1130 AGGTGCCCCGCTGAGCTCTCTCATCTGCTGAGCCCTGACCTGACGAGACGTGTGGGG 1189  
Db 4307 AGGTGCCCCGCTGAGCTCTCTCATCTGCTGAGCCCTGACCTGACGAGACGTGTGGGG 4366  
Qy 1190 AGAGGTGCCCCCTTCAACCAACGCTCAAGCTGCAAGCGCTTTTGTGTTATC 1249  
Db 4367 AGAGGTGCCCCCTTCAACCAACGCTCAAGCTGCAAGCGCTTTTGTGTTATC 4426  
Qy 1250 AGCTACCCCTGAGAGGCTGAGCGGCTGCTCTCAATCTGAGAGAAAGTGTGCTGA 1309  
Db 4427 AGCTACCCCTGAGAGGCTGAGCGGCTGCTCTCAATCTGAGAGAAAGTGTGCTGA 4486  
Qy 1310 AGCGCTTCTGCGATGAGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTGTGAGT 1369  
Db 4487 AGCGCTTCTGCGATGAGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTGTGAGT 4546  
Qy 1370 ACGACCTATAGTTCAGAGACTCGGGAGTCTGAGCTCGTCTTTCATGAAACA 1429  
Db 4547 ACGACCTATAGTTCAGAGACTCGGGAGTCTGAGCTCGTCTTTCATGAAACA 4606  
Qy 1430 TCCCTTCTCCAGACGCACTGTGAGAGTCAACCTTCTTCCACTGCTCCAGG 1489  
Db 4607 TCCCTTCTCCAGACGCACTGTGAGAGTCAACCTTCTTCCACTGCTCCAGG 4666  
Qy 1490 ACTTCTGTGCGGCTTGTACTACGTGTGAGAGGCTGGAATCGAGCCTGCTGCC 1549  
Db 4667 ACTTCTGTGCGGCTTGTACTACGTGTGAGAGGCTGGAATCGAGCCTGCTGCC 4726  
Qy 1550 CTCTGTACGTGAGAGCAAAAGAGTCTCATGAGCTTAAACAGGACGCTTCCATATC 1609  
Db 4727 CTCTGTACGTGAGAGCAAAAGAGTCTCATGAGCTTAAACAGGACGCTTCCATATC 4786  
Qy 1610 ACTGCTTGTGATGAGAGGTTTCTTGTGCTGTGAGAGAGAGAGAGAGAGAGAGAG 1669  
Db 4787 ACTGCTTGTGATGAGAGGTTTCTTGTGCTGTGAGAGAGAGAGAGAGAGAGAGAG 4846  
Qy 1670 TGAAGGTCTGCTGAGCTGTGCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 1729  
Db 4847 TGAAGGTCTGCTGAGCTGTGCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 4906  
Qy 1730 TCTCTCTGTGAGAGAGCTTAAATGCAACCCAGAGAGACCTCTGACGCTTCC 1789  
Db 4907 TCTCTCTGTGAGAGAGCTTAAATGCAACCCAGAGAGACCTCTGACGCTTCC 4966  
Qy 1790 ACTGCTTTTGAAGCTCAAGACAAAGAGTTGTGCTGTGAGAGAGAGAGAGAGAGAG 1849  
Db 4967 ACTGCTTTTGAAGCTCAAGACAAAGAGTTGTGCTGTGAGAGAGAGAGAGAGAGAG 5026  
Qy 1850 AAGTGTGCTTCCGATTAACAGAACTGAGCTTGATGACATCTTCTGCTCCAGC 1909

Db 5027 AAGTGTGCTTCCGATTAACAGAACTGAGCTTGATGACATCTTCTGCTCCAGC 5086  
Qy 1910 ACTGTCCGTATTTGCGGAAATTTGCGGTGATGTCAAGAGGATTTTCCAGAGATAGT 1969  
Db 5087 ACTGTCCGTATTTGCGGAAATTTGCGGTGATGTCAAGAGGATTTTCCAGAGATAGT 5146  
Qy 1970 CCGCTGAGGCAATGCTGTGTGCTCCCTTATGATGAGCGGATTAAGACCTCATTTAGAGAGC 2029  
Db 5147 CCGCTGAGGCAATGCTGTGTGCTCCCTTATGATGAGCGGATTAAGACCTCATTTAGAGAGC 5206  
Qy 2030 AGTGGAAAGATTTCTGCTCCATGTTGAGCAACCAACCTGCGGACGTGACCTGG 2089  
Db 5207 AGTGGAAAGATTTCTGCTCCATGTTGAGCAACCAACCTGCGGACGTGACCTGG 5266  
Qy 2090 GCACACGATCTGACAGAGCGGGCAATGAAGACCTGTGTGCCAGCTGAGGATCCCA 2149  
Db 5267 GCACACGATCTGACAGAGCGGGCAATGAAGACCTGTGTGCCAGCTGAGGATCCCA 5326  
Qy 2150 CCTGCAAGATPACAGACCTGATGTTTGAAGATGACAGATTAACCTGTGTGAGAGACC 2209  
Db 5327 CCTGCAAGATPACAGACCTGATGTTTGAAGATGACAGATTAACCTGTGTGAGAGACC 5386  
Qy 2210 TCTGAGAAATGTCATGAGCCCAACCGTAACTTAAGATCTTCAACTTGGAGGACCAAC 2269  
Db 5387 TCTGAGAAATGTCATGAGCCCAACCGTAACTTAAGATCTTCAACTTGGAGGACCAAC 5446  
Qy 2270 TGAAGAAAGAAATGTAAGATGAGCGGTGTGAAGCTTTAAACACCCAAATGTTGTGG 2329  
Db 5447 TGAAGAAAGAAATGTAAGATGAGCGGTGTGAAGCTTTAAACACCCAAATGTTGTGG 5506  
Qy 2330 AGTCTTTGAGGCTGATTTGCTGTGATGACCCCATGCTGTAACTGAAGATCTCCCAA 2389  
Db 5507 AGTCTTTGAGGCTGATTTGCTGTGATGACCCCATGCTGTAACTGAAGATCTCCCAA 5566  
Qy 2390 TCTTTAGACTCTCCCAAGCTTGAATCTGTAGCTTGGGAGAAACAAGTGAACAACC 2449  
Db 5567 TCTTTAGACTCTCCCAAGCTTGAATCTGTAGCTTGGGAGAAACAAGTGAACAACC 5626  
Qy 2450 AGGAGATTAAGCTCTGATGATGAGCTTGAAGATCTCCAGTGGGCTGTGAGAAAGCTGA 2509  
Db 5627 AGGAGATTAAGCTCTGATGATGAGCTTGAAGATCTCCAGTGGGCTGTGAGAAAGCTGA 5686  
Qy 2510 TACTGAGGACTGTGAGATCAACAGCAGGATTCAGAGCTGTGAGCTGAGCTGCTCA 2569  
Db 5687 TACTGAGGACTGTGAGATCAACAGCAGGATTCAGAGCTGTGAGCTGAGCTGCTCA 5746  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTTATCAACAACAGCTTGGGAAACGAGCTGTA 2629  
Db 5747 GCAACCGGAGCTTGAACAACCTGTGCTTATCAACAACAGCTTGGGAAACGAGCTGTA 5806  
Qy 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTATCTGCAAGAGCTGATGCTGAATC 2689  
Db 5807 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTATCTGCAAGAGCTGATGCTGAATC 5866  
Qy 2690 AGTGCACCTGGAACAGGCTGTGTGTGTTTCTTGAACCTTGTGAGTATGAGTAACTCAT 2749  
Db 5867 AGTGCACCTGGAACAGGCTGTGTGTGTTTCTTGAACCTTGTGAGTATGAGTAACTCAT 5926  
Qy 2750 GGTGACGCACTGAGCTTATGACATGAACCTGTGAAAGACATGAGCTGTAAGCTTGT 2809  
Db 5927 GGTGACGCACTGAGCTTATGACATGAACCTGTGAAAGACATGAGCTGTAAGCTTGT 5986  
Qy 2810 GCGAGGTCAATGAGAGAACATCTTGTATCTCCAGAGACCTGAGATTGTGTAAGTGTATC 2869  
Db 5987 GCGAGGTCAATGAGAGAACATCTTGTATCTCCAGAGACCTGAGATTGTGTAAGTGTATC 6046  
Qy 2870 TCACCGCGAGTGTGAGAGAGTCTGTCTGTGTGATCTGAGAGAGAGACACCTGAAGA 2929  
Db 6047 TCACCGCGAGTGTGAGAGAGTCTGTCTGTGTGATCTGAGAGAGAGACACCTGAAGA 6106  
Qy 2930 GCTGTGATCTCAACGACAAATGCTGTGGTGAACGCTGTGGGTTGCTGACCTGTGAGAGAC 2989  
Db 6107 GCTGTGATCTCAACGACAAATGCTGTGGTGAACGCTGTGGGTTGCTGACCTGTGAGAGAC 6166

Accession	Gene	Species	Length (bp)	GC Content (%)	GC Skew	GC Bias	GC Bias2	GC Bias3	GC Bias4	GC Bias5	GC Bias6	GC Bias7	GC Bias8	GC Bias9	GC Bias10	GC Bias11	GC Bias12	GC Bias13	GC Bias14	GC Bias15	GC Bias16	GC Bias17	GC Bias18	GC Bias19	GC Bias20	GC Bias21	GC Bias22	GC Bias23	GC Bias24	GC Bias25	GC Bias26	GC Bias27	GC Bias28	GC Bias29	GC Bias30	GC Bias31	GC Bias32	GC Bias33	GC Bias34	GC Bias35	GC Bias36	GC Bias37	GC Bias38	GC Bias39	GC Bias40	GC Bias41	GC Bias42	GC Bias43	GC Bias44	GC Bias45	GC Bias46	GC Bias47	GC Bias48	GC Bias49	GC Bias50	GC Bias51	GC Bias52	GC Bias53	GC Bias54	GC Bias55	GC Bias56	GC Bias57	GC Bias58	GC Bias59	GC Bias60	GC Bias61	GC Bias62	GC Bias63	GC Bias64	GC Bias65	GC Bias66	GC Bias67	GC Bias68	GC Bias69	GC Bias70	GC Bias71	GC Bias72	GC Bias73	GC Bias74	GC Bias75	GC Bias76	GC Bias77	GC Bias78	GC Bias79	GC Bias80	GC Bias81	GC Bias82	GC Bias83	GC Bias84	GC Bias85	GC Bias86	GC Bias87	GC Bias88	GC Bias89	GC Bias90	GC Bias91	GC Bias92	GC Bias93	GC Bias94	GC Bias95	GC Bias96	GC Bias97	GC Bias98	GC Bias99	GC Bias100
2990	TGAAGCAAAAGACAGCTTCGACGACACTCGGTTGAAGGCATGTGGAAGCTTCTG	3043	6167	TGAAGCAAAAGACAGCTTCGACGACACTCGGTTGAAGGCATGTGGAAGCTTCTG	6226	3050	ATTGCTGTGAGCACTCTCTTGGCCCTTCTCGACACGGGCATCTGACCACTTAAAC	3109	6227	ATTGCTGTGAGCACTCTCTTGGCCCTTCTCGACACGGGCATCTGACCACTTAAAC	6286	3110	TGCTGTCAGATTAATCTTCACTGCTCCCAAGGAATGATGAGCTGTGCTGGCCTTGTCTGTC	3169	6287	TGCTGTCAGATTAATCTTCACTGCTCCCAAGGAATGATGAGCTGTGCTGGCCTTGTCTGTC	6346	3170	CCAGCTTAACCTTACAGATAATTGG	3194	6347	CCAGCTTAACCTTACAGATAATTGG	6371																																																																																		

|||||  
Db 309 GATGTACGTCTAGTTTGAACAACATGCTGTGCTGATCGGCGGAAATGCAAACTGTGGCT 368  
|||  
Qy 526 GGTCTTTTGAATTGACACCGGTGGGGCTTCCGGCTCTGACAGGTGTTTCTGACAGGAAG 585  
|||  
Db 369 GGTCTTTTGAATTGACACCGGTGGGGCTTCCGGCTCTGACAGGTGTTTCTGACAGGAAG 428  
|||  
Qy 586 TCAGGAATTTGGGAAATCGGCTCTAGCCAGAGGATCGTGTCTGTGGCCGAAGTGA 645  
|||  
Db 429 TCAGGAATTTGGGAAATCGGCTCTAGCCAGAGGATCGTGTCTGTGGCCGAAGTGA 488  
|||  
Qy 646 CTCTACCGAGGAATGTTCTCTAGTCTTCTTCTCCCGTTTAAAGATCGACGGGAAG 705  
|||  
Db 489 CTCTACCGAGGAATGTTCTCTAGTCTTCTTCTCCCGTTTAAAGATCGACGGGAAG 548  
|||  
Qy 706 AAGGAGACAGTGTCAAGAGTTGATCTCCAGGGGTGGCAGACTCCAGGCTCCGGTG 765  
|||  
Db 549 AAGGAGACAGTGTCAAGAGTTGATCTCCAGGGGTGGCAGACTCCAGGCTCCGGTG 608  
|||  
Qy 766 ACGGAGATCATGTCCCGACAGAAAGCTGTTGTTCAATCATTTGACGTTTGCATGACCTG 825  
|||  
Db 609 ACGGAGATCATGTCCCGACAGAAAGCTGTTGTTCAATCATTTGACGTTTGCATGACCTG 668  
|||  
Qy 826 GGTCTGTCTCAACATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGACGCTTCG 885  
|||  
Db 669 GGTCTGTCTCAACATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGACGCTTCG 728  
|||  
Qy 886 TTCAACCTCATACCAAGTCTGTGAGGAAAGTCCGTCCCGTGAAGTCCCTGATGCTC 945  
|||  
Db 729 TTCAACCTCATACCAAGTCTGTGAGGAAAGTCCGTCCCGTGAAGTCCCTGATGCTC 788  
|||  
Qy 946 ACCCTCAGAGACGTGGGACAGAGAGCTCAAGTCAAGATCGTGTCTCCCGTTTACCTG 1005  
|||  
Db 789 ACCCTCAGAGACGTGGGACAGAGAGCTCAAGTCAAGATCGTGTCTCCCGTTTACCTG 848  
|||  
Qy 1006 TTACTTAAAGGAATCTCCGGGAAACAAAGATCCAATTGCTCTTGAACGCGGGATTGGT 1065  
|||  
Db 849 TTACTTAAAGGAATCTCCGGGAAACAAAGATCCAATTGCTCTTGAAGCGGGATTGGT 908  
|||  
Qy 1066 GAGATCAAGAGACAAAGGTTGCTGCGATCAATGAAACAACCTGAGCTCTGACAG 1125  
|||  
Db 909 GAGATCAAGAGACAAAGGTTGCTGCGATCAATGAAACAACCTGAGCTCTGACAG 968  
|||  
Qy 1126 TGCCAGGTGCGCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGAGTGG 1185  
|||  
Db 969 TGCCAGGTGCGCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGAGTGG 1028  
|||  
Qy 1186 GGGGAGAGCTGCGCCCTTCAACCAACGCTCAAGGCTGCAAGCCGCTTTTGTGTTT 1245  
|||  
Db 1029 GGGGAGAGCTGCGCCCTTCAACCAACGCTCAAGGCTGCAAGCCGCTTTTGTGTTT 1088  
|||  
Qy 1246 CATCAGCTCAACCCCTCGAGGGGTGGTCCGGGCTGTCTCAATCTGAGGAGAAAGTTGTC 1305  
|||  
Db 1089 CATCAGCTCAACCCCTCGAGGGGTGGTCCGGGCTGTCTCAATCTGAGGAGAAAGTTGTC 1148  
|||  
Qy 1306 CTGAAGCGCTTCTCGCGTATGCTGTGAGGAGATGAGGAAATAGGAAGTCAGTGTTTGA 1365  
|||  
Db 1149 CTGAAGCGCTTCTCGCGTATGCTGTGAGGAGATGAGGAAATAGGAAGTCAGTGTTTGA 1208  
|||  
Qy 1366 GGTGACGACTTCATGATTCAGAGACTCGGGAGTCTGAGCTCCGTCTGTGTTTCAATG 1425  
|||  
Db 1209 GGTGACGACTTCATGATTCAGAGACTCGGGAGTCTGAGCTCCGTCTGTGTTTCAATG 1268  
|||  
Qy 1426 AACTCCTTCTCCAGACAGCACTGTGAGAGATCAACCTTCTTCCACTGATGCTC 1485  
|||  
Db 1269 AACTCCTTCTCTCCAGACAGCACTGTGAGAGATCAACCTTCTTCCACTGATGCTC 1338  
|||  
Qy 1486 CAGACCTTCTGTGCGCTTGTACTACGTGTTAGAGGGCTGGAATGACCGACGCTCTC 1545  
|||  
Db 1329 CAGACCTTCTGTGCGCTTGTACTACGTGTTAGAGGGCTGGAATGACCGACGCTCTC 1388  
|||  
Qy 1546 TGCCCTCTGTACGTTGAGAGACAAAGAGTCAATGAGCTTAAACAGGAGCTTCAT 1605  
|||

Db 1389 TGCCCTCTGTACGTTGAGAGACAAAGAGTCCATGAGGCTTAAACAGCAGCTTCAT 1448  
|||  
Qy 1606 ATCCACTCGCTTTGAGTGAAGCGTTTCTGTTTGGCTCTGTGAGCGAAGAGTAAAGAG 1665  
|||  
Db 1449 ATCCACTCGCTTTGAGTGAAGCGTTTCTGTTTGGCTCTGTGAGCGAAGAGTAAAGAG 1508  
|||  
Qy 1666 CCACTGAGAGTCTGCTGGGCTGTCCGTTTCCCTGGGGGTGAAGCAAGACTTCTGCAC 1725  
|||  
Db 1509 CCACTGAGAGTCTGCTGGGCTGTCCGTTTCCCTGGGGGTGAAGCAAGACTTCTGCAC 1568  
|||  
Qy 1726 TGGGTCTCTGTGTGGGTCAAGACCTTAATGCCACCAACCCAGAGACACCTTGGAGGCC 1785  
|||  
Db 1569 TGGGTCTCTGTGTGGGTCAAGACCTTAATGCCACCAACCCAGAGACACCTTGGAGGCC 1628  
|||  
Qy 1786 TTCCACTGTCTTTTGCAGACTCAAGACAAAGATTTTGTGCTTGGATTTAAACAGCTTC 1845  
|||  
Db 1629 TTCCACTGTCTTTTGCAGACTCAAGACAAAGATTTTGTGCTTGGATTTAAACAGCTTC 1688  
|||  
Qy 1846 CAAGAGTGTGGCTTCCGATTTAACAGAACCTTGACCTTGATAGCATCTTCTGTGCTC 1905  
|||  
Db 1689 CAAGAGTGTGGCTTCCGATTTAACAGAACCTTGACCTTGATAGCATCTTCTGTGCTC 1748  
|||  
Qy 1906 CAGCACTGTCCGTATTTTGGGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAT 1965  
|||  
Db 1749 CAGCACTGTCCGTATTTTGGGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAT 1808  
|||  
Qy 1966 GAGTCCGCTGAGGAGATGCTGTGTGCTCTATGATGAGGAGATAAGACCTTCATTTGAG 2025  
|||  
Db 1809 GAGTCCGCTGAGGAGATGCTGTGTGCTCTATGATGAGGAGATAAGACCTTCATTTGAG 1868  
|||  
Qy 2026 GAGCAGTGGAGAGATTTCTGCTCCATGCTTGGCAACCCACACCTGCGGACGTGAC 2085  
|||  
Db 1869 GAGCAGTGGAGAGATTTCTGCTCATGTCTTGGCAACCCACACCTGCGGACGTGAC 1928  
|||  
Qy 2086 CTGGGACAGCAGATCTCTGACAGAGCGGCCATGAAGACCTGTGTGCCAAGCTGAGCAT 2145  
|||  
Db 1929 CTGGGACAGCAGATCTCTGACAGAGCGGCCATGAAGACCTGTGTGCCAAGCTGAGCAT 1988  
|||  
Qy 2146 CCCACTGCAAGATACAGACCCCTGATGTTTGAAGATGACAGATTAACCCCTGGTGTGAG 2205  
|||  
Db 1989 CCCACTGCAAGATACAGACCCCTGATGTTTGAAGATGACAGATTAACCCCTGGTGTGAG 2048  
|||  
Qy 2206 CACCTTGAAGATCGTCAATGCGCAACCGTAAACCTTAAGATCCCTCAACCTTGGAGGAC 2265  
|||  
Db 2049 CACCTTGAAGATCGTCAATGCGCAACCGTAAACCTTAAGATCCCTCAACCTTGGAGGAC 2108  
|||  
Qy 2266 CACTGAAGAGAGAGATGTAAGATGCGGTGTACAGCTTTAAACACCCAAAATGTTTG 2325  
|||  
Db 2109 CACTGAAGAGAGAGATGTAAGATGCGGTGTACAGCTTTAAACACCCAAAATGTTTG 2168  
|||  
Qy 2326 TTGAGATCTTTGAGGCTGGAATGCTGTGATTTGACCCATGCGTGTACTGAAGATCTCC 2385  
|||  
Db 2169 TTGAGATCTTTGAGGCTGGAATGCTGTGATTTGACCCATGCGTGTACTGAAGATCTCC 2228  
|||  
Qy 2386 CAAATCTTACGACCTTCCCGACGCTGAAATCTCTGAGCTGTGAGGAAACAAAGTGA 2445  
|||  
Db 2229 CAAATCTTACGACCTTCCCGACGCTGAAATCTCTGAGCTGTGAGGAAACAAAGTGA 2288  
|||  
Qy 2446 GACCAAGGAGATATGCTCTCAATGATGCTTGAAGATCTCCAGATGCGCTTGCAGAG 2505  
|||  
Db 2289 GACCAAGGAGATATGCTCTCAATGATGCTTGAAGATCTCCAGATGCGCTTGCAGAG 2348  
|||  
Qy 2506 CTGATACGTGAGGACTGTGGATCAAGCAACGAGGTTGCCAGAGTGTGAGCTCAGCCCTC 2565  
|||  
Db 2349 CTGATACGTGAGGACTGTGGATCAAGCAACGAGGTTGCCAGAGTGTGAGCTCAGCCCTC 2408  
|||  
Qy 2566 GTCAGCAACCGGAGCTTGACACACTGTGTCTATCCAAACAACGCTTGGGAAACGAAGT 2625  
|||  
Db 2409 GTCAGCAACCGGAGCTTGACACACTGTGTCTATCCAAACAACGCTTGGGAAACGAAGT 2468  
|||  
Qy 2626 GTAAATCTACTGTGTGCATCAATGAGGCTTCCCACTGTAGTGTGACAGAGCTGATCTG 2685  
|||  
Db 2469 GTAAATCTACTGTGTGCATCAATGAGGCTTCCCACTGTAGTGTGACAGAGCTGATCTG 2528  
|||



QY 2686 AATGACGTCACCTGAGACGCGCTGGCTGCTGCTTCTTGACCTTGGCTTAAAGGTAAC 2745  
DB 2529 AATGACGTCACCTGAGACGCGCTGGCTGCTGCTTCTTGACCTTGGCTTAAAGGTAAC 2588  
QY 2746 TCATGCTGACGACCTGAGACCTTGAACATGAACCTTGGAGAACATGGCTGAGCTT 2805  
DB 2589 TCATGCTGACGACCTGAGACCTTGAACATGAACCTTGGAGAACATGGCTGAGCTT 2648  
QY 2806 CTGTGCGAGCTCATGAGAGAACATTTGTCTCATCTCCAGACCTGAGATTGGTAAAGTGT 2865  
DB 2649 CTGTGCGAGCTCATGAGAGAACATTTGTCTCATCTCCAGACCTGAGATTGGTAAAGTGT 2708  
QY 2866 CATCTCACCGCGCGCTGTGAGAGCTGTCTCTGTGTATCTCGAGAGACGACGACCTG 2925  
DB 2709 CATCTCACCGCGCGCTGTGAGAGCTGTCTCTGTGTATCTCGAGAGACGACGACCTG 2768  
QY 2926 AAGAGCTGTGATCTCAACGACATGCTGCTGGTGAAGCTGCTGCTGCTGCTGCTG 2985  
DB 2769 AAGAGCTGTGATCTCAACGACATGCTGCTGGTGAAGCTGCTGCTGCTGCTGCTG 2828  
QY 2986 GGAATGAGCAAAAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCAATGTGACTACT 3045  
DB 2829 GGAATGAGCAAAAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCAATGTGACTACT 2888  
QY 3046 TCTGATTTGTGAGGCACTCTCTTGGGCTTCTGACGACCGGATCTGACCACTCTA 3105  
DB 2889 TCTGATTTGTGAGGCACTCTCTTGGGCTTCTGACGACCGGATCTGACCACTCTA 2948  
QY 3106 AACCTGTGACAAATTAATCTTCACTGCTCCCAAGAAATGATGATGCTGCTGCTTGGC 3165  
DB 2949 AACCTGTGACAAATTAATCTTCACTGCTCCCAAGAAATGATGATGCTGCTGCTTGGC 3008  
QY 3166 TGTCCCAAGCTTAATTAATTAATTTGGCTGTGAGAAATGACAGTCTTGGCAATA 3225  
DB 3009 TGTCCCAAGCTTAATTAATTAATTTGGCTGTGAGAAATGACAGTCTTGGCAATA 3068  
QY 3226 AGGAGCTGTGAGGAGAGTGAAGTCACTCAAGCCCGGCTGATTAACGCTAGTTGG 3285  
DB 3069 AGGAGCTGTGAGGAGAGTGAAGTCACTCAAGCCCGGCTGATTAACGCTAGTTGG 3128  
QY 3286 CATCTTTTGTGATGAGTACCG 3308  
DB 3129 CATCTTTTGTGATGAGTACCG 3151

RESULT 10  
ABK48628  
ID ABK48628 standard; cDNA; 3900 BP.  
XX  
AC ABK48628;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Human MATER cDNA.  
XX  
KW Human; gene; ss; contraceptive; antiinfertility; MATER;  
KW maternal antigen that embryos require; MATER null phenotype; oocyte;  
KW early embryonic survival; premature ovarian failure; POF;  
KW autoimmune infertility; chromosome 19; gene therapy; fertility.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..3603  
XX FT /\*tag= a  
XX FT /product= "Human MATER"  
XX  
XX MO200232955-A1.  
XX 25-APR-2002.  
XX  
XX 04-APR-2001; 2001WO-US010981.  
XX

XX  
PR 18-OCT-2000; 2000US-0241510P.  
XX  
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nelson LM, Tong Z;  
XX  
DR WPI: 2002-454595/48.  
XX  
DR P-PSDB; AAU79526.  
XX  
PT New isolated human Maternal Antigen That Embryos Require protein and  
PT polynucleotide, useful in diagnosing or treating fertility or reduced  
PT fertility, or as a contraceptive.  
PS  
PS Claim 11: Page 80-85; 93pp; English.  
XX  
XX The invention discloses an isolated human MATER (Maternal Antigen That  
XX Embryos Require) protein, which can complement a MATER null phenotype in  
XX which zygotes arising from the oocyte do not progress beyond the two-cell  
XX stage. MATER is required for early embryonic survival and abnormal levels  
XX of the protein can lead to premature ovarian failure (POF) and can be  
XX caused by under or over expression of MATER or an autoimmune response to  
XX MATER. MATER is a single-copy maternal effect gene found on chromosome  
XX 19. The MATER protein and polynucleotide, by gene therapy, are useful in  
XX diagnosing or treating fertility and reduced fertility. In particular,  
XX the MATER protein is useful as a contraceptive agent, or for influencing  
XX (either inhibiting or enhancing) fertility and can be used to detect a  
XX predisposition to infertility or reduced fertility, or for presymptomatic  
XX screening of an individual for infertility/reduced fertility. The protein  
XX and polynucleotide are also useful for detecting an excess or deficiency,  
XX or genetic mutation, of the MATER protein in a mammalian subject (e.g. a  
XX human or a mouse) or for screening for a compound useful in influencing  
XX MATER-mediated fertility. The sequence presented is the human MATER cDNA.  
XX Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to  
XX published human genomic DNA sequences, while the remainder of the  
XX sequence was determined by direct cloning of human ovarian cDNAs  
XX  
SQ Sequence: 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;  
Query Match 67.2%; Score 2713; DB 6; Length 3900;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 290 AAATTTCACAGCTATGGAACAAGAGTGGCCACAGCAGCAGACAGAGAACAGAGAA 349  
DB 566 AAATTTCACAGCTATGGAACAAGAGTGGCCACAGCAGCAGACAGAGAACAGAGAA 625  
QY 350 TTTCAAGCTATGGAACAAGAGTGGCCACAGCAGCAGACAGAGAACAGAGACAATG 409  
DB 626 TTTCAAGCTATGGAACAAGAGTGGCCACAGCAGCAGACAGAGAACAGAGACAATG 685  
QY 410 GAGGTGACACATGGGACTATCAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGATG 469  
DB 686 GAGGTGACACATGGGACTATCAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGATG 745  
QY 470 TAGCTGTAGTTTGAAGAACTGCTGCTGACCTGGCCGGAATGCAAACTTTGGCTG 529  
DB 746 TAGCTGTAGTTTGAAGAACTGCTGCTGACCTGGCCGGAATGCAAACTTTGGCTG 805  
QY 530 CTTTGTATTCACACCGGTGGGCTTCCGCTCGACAGCTGTGTTGACAGGAAAGTCA 589  
DB 806 CTTTGTATTCACACCGGTGGGCTTCCGCTCGACAGCTGTGTTGACAGGAAAGTCA 865  
QY 590 GAATTGGGAATTCGCTTACCCAGAGAGATGCTGTGCTGGGCGCAAGTGTGACTCT 649  
DB 866 GAATTGGGAATTCGCTTACCCAGAGAGATGCTGTGCTGGGCGCAAGTGTGACTCT 925  
QY 650 ACCAGGGAATGTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACGGAAGAG 709  
DB 926 ACCAGGGAATGTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACGGAAGAG 985  
QY 710 AGAGAGTGTACAGAGTATCTCCAGGAGTGGCCAGACTCCAGCTCCGCTGACGG 769

Db 986 AGACAGTGTCA CAGAGTTCTCTCCAGGAGTGGCCAGACTCCCAAGCTCCGGTGAACG 1045  
Qy 770 AGATCATGTCCCGACCAAGAAAGGCTGTGTTCATCATTTGACGTTTGCATGACCTGGGCT 829  
Db 1046 AGATCATGTCCCGACCAAGAAAGGCTGTGTTCATCATTTGACGTTTGCATGACCTGGGCT 1105  
Qy 830 CTGTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGAGAGCTCCGTTCA 889  
Db 1106 CTGTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGAGAGCTCCGTTCA 1165  
Qy 890 CCTCTATACGAGTGTGCTGAGAGAGGCTCTGCTCCCTGAGTCTTCTGATTCGTCAACG 949  
Db 1166 CCTCTATACGAGTGTGCTGAGAGAGGCTCTGCTCCCTGAGTCTTCTGATTCGTCAACG 1225  
Qy 950 TCAGAGAGTGGGACAGAGAGGCTCAAGTCAGAGGTCGTGCTCCCGCTTACCTGTTAG 1009  
Db 1226 TCAGAGAGTGGGACAGAGAGGCTCAAGTCAGAGGTCGTGCTCCCGCTTACCTGTTAG 1285  
Qy 1010 TTAGAGGATCTCCGGGGACAAAGAAATCACTTGCTCTTGAAGCGCGGAAATTGGTGAAC 1069  
Db 1286 TTAGAGGATCTCCGGGGACAAAGAAATCACTTGCTCTTGAAGCGCGGAAATTGGTGAAC 1345  
Qy 1070 ATCAGAGACACAAAGGTTGGGTGCTGATCATGAAACCTGAGCTGCTCCACAGTACC 1129  
Db 1346 ATCAGAGACACAAAGGTTGGGTGCTGATCATCMAACCTGAGCTGCTCCACAGTACC 1405  
Qy 1130 AGGTGCCGCGGGGCTCTCTCATCTGCGGCTCTGACAGCTGACAGGACGTGGTGGGG 1189  
Db 1406 AGGTGCCGCGGGGCTCTCTCATCTGCGGCTCTGACAGCTGACAGGACGTGGTGGGG 1465  
Qy 1190 AGAGCGTGCCTCCCTTCAACAAACGCTCAGAGGCTGACAGCGCTTTTGTGTTATC 1249  
Db 1466 AGAGCGTGCCTCCCTTCAACAAACGCTCAGAGGCTGACAGCGCTTTTGTGTTATC 1525  
Qy 1250 AGCTCAACCTCTGAGAGGCTGCTCGGCGCTGTCTCAATCTGAGAGAAAGTGTCTGCA 1309  
Db 1526 AGCTCAACCTCTGAGAGGCTGCTCGGCGCTGTCTCAATCTGAGAGAAAGTGTCTGCA 1585  
Qy 1310 AGCGCTTCTGCGATGAGCTGAGAGGAGTGTGAATAGGAAGTCAGTGTGTAAGTGTG 1369  
Db 1586 AGCGCTTCTGCGATGAGCTGAGAGGAGTGTGAATAGGAAGTCAGTGTGTAAGTGTG 1645  
Qy 1370 ACGACCTCATGATTCMAAGACTCGGAGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1429  
Db 1646 ACGACCTCATGATTCMAAGACTCGGAGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1705  
Qy 1430 TCTCTTCTCCAGACAGCTGTGAGAGTACTACCTTCTTCACTCACTGCTCCAGG 1489  
Db 1706 TCTCTTCTCCAGACAGCTGTGAGAGTACTACCTTCTTCACTCACTGCTCCAGG 1765  
Qy 1490 ACTTCTGCGCGCTTGTACTACGTTGAGAGGCTGGAATGGAAGTCAGCTCTGCTG 1549  
Db 1766 ACTTCTGCGCGCTTGTACTACGTTGAGAGGCTGGAATGGAAGTCAGCTCTGCTG 1825  
Qy 1550 CTCTGTACGTTGAGAGCAAAAGAGTCAATGAGACTTAAACAGGCAAGCTTCCATATCC 1609  
Db 1826 CTCTGTACGTTGAGAGCAAAAGAGTCAATGAGACTTAAACAGGCAAGCTTCCATATCC 1885  
Qy 1610 ACTGCTTTGAGATGAAGGCTTCTGTTGGCTCTGTAGAGGAGAGCTAAGAGGCGAC 1669  
Db 1886 ACTGCTTTGAGATGAAGGCTTCTGTTGGCTCTGTAGAGGAGAGCTAAGAGGCGAC 1945  
Qy 1670 TGAAGGTCCTGCTGGGCTGTCCCGTTCCCTGAGGAGTGAAGAGGCTTCTGCACTGG 1729  
Db 1946 TGAAGGTCCTGCTGGGCTGTCCCGTTCCCGTTGAGGAGTGAAGAGGCTTCTGCACTGG 2005  
Qy 1730 TCTCTGTGTTGAGTCAAGAGCTATATGCAACCCCAAGAGACACCTGAGCGCTTCC 1789  
Db 2006 TCTCTGTGTTGAGTCAAGAGCTATATGCAACCCCAAGAGACACCTGAGCGCTTCC 2065  
Qy 1790 ACTGTCTTTTGAAGCTCAAGACAAAGTGTGCTGCTGGCATTTAAAGCTTCCAG 1849  
Db 2066 ACTGTCTTTTGAAGCTCAAGACAAAGTGTGCTGCTGGCATTTAAAGCTTCCAG 2125

Qy 1850 AAGTGTGCTTCCGATTAACAGAACCTTGAATCTGTATAGCATCTTCTGCTGCTCAGC 1909  
Db 2126 AAGTGTGCTTCCGATTAACAGAACCTTGAATCTGTATAGCATCTTCTGCTGCTCAGC 2185  
Qy 1910 ACTGTCCGATTTTGGGAAAATTTGGGTGATGTCAAAAGGATCTTCCCAAGATGAGT 1969  
Db 2186 ACTGTCCGATTTTGGGAAAATTTGGGTGATGTCAAAAGGATCTTCCCAAGATGAGT 2245  
Qy 1970 CCGCTGAGGAGATGTCCTGAGTCCCTGTATAGATGCGGGAATTAAGACCTCATTTAGAGAGC 2029  
Db 2246 CCGCTGAGGAGATGTCCTGAGTCCCTGTATAGATGCGGGAATTAAGACCTCATTTAGAGAGC 2305  
Qy 2030 AGTGGAAAGATTTCTGCTCCATGCTTGAACCCCAACCAACCTGAGAGCTGAGCTGG 2089  
Db 2306 AGTGGAAAGATTTCTGCTCCATGCTTGAACCCCAACCAACCTGAGAGCTGAGCTGG 2365  
Qy 2090 GCAGCAGATCTTGAACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGATCCCA 2149  
Db 2366 GCAGCAGATCTTGAACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGATCCCA 2425  
Qy 2150 CCTGCAAGATTCAGACCTGATGTTTGAAGATGACAGATTACCCTGCTGTGAGCAGC 2209  
Db 2426 CCTGCAAGATTCAGACCTGATGTTTGAAGATGACAGATTACCCTGCTGTGAGCAGC 2485  
Qy 2210 TCTGAGATTCGTATGAGCCCAACCGTAACCTTAAGTCCCTCACTTGGAGGACCCAC 2269  
Db 2486 TCTGAGATTCGTATGAGCCCAACCGTAACCTTAAGTCCCTCACTTGGAGGACCCAC 2545  
Qy 2270 TGAAGAAAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2546 TGAAGAAAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2605  
Qy 2230 AGTCTTTGAGGCTGAGATTTGCTGTGAGTTGACCCAATGCTGTACTGAAGATCTCCCAA 2389  
Db 2606 AGTCTTTGAGGCTGAGATTTGCTGTGAGTTGACCCAATGCTGTACTGAAGATCTCCCAA 2665  
Qy 2290 TCTTTAGACCTCCCGACGCTGAAATCTCTGAGCTTGGCAGGAAACAAAGTGAACAAC 2449  
Db 2666 TCTTTAGACCTCCCGACGCTGAAATCTCTGAGCTTGGCAGGAAACAAAGTGAACAAC 2725  
Qy 2450 AGGAGTATATCTCTGATGATGCTCTTGAAGATCTCCAGTGGCTCTGAGAACTGCA 2509  
Db 2726 AGGAGTATATCTCTGATGATGCTCTTGAAGATCTCCAGTGGCTCTGAGAACTGCA 2785  
Qy 2510 TACTGAGAGACTGAGATCAACGCCAGGTTGCAAGATCTGAGCTCCGCTGCA 2569  
Db 2786 TACTGAGAGACTGAGATCAACGCCAGGTTGCAAGATCTGAGCTCCGCTGCA 2845  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTATCAACAACAGCTGGGAAACGAAGTGTAA 2629  
Db 2846 GCAACCGGAGCTTGAACAACCTGTGCTATCAACAACAGCTGGGAAACGAAGTGTAA 2905  
Qy 2630 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTATCTGAGAGCTGATCTGATC 2689  
Db 2906 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTATCTGAGAGCTGATCTGATC 2965  
Qy 2690 AGTCCACCTGGAACAGGCTGAGTGTGTTCTTGAACCTTGAAGCTTAAGGTAATCAT 2749  
Db 2966 AGTCCACCTGGAACAGGCTGAGTGTGTTCTTGAACCTTGAAGCTTAAGGTAATCAT 3025  
Qy 2750 GGTGAGCAGCTGAGCTTGAATGAATGAACCTGTGGAAGACAAATGAGCTGAGCTTCTGT 2809  
Db 3026 GGTGAGCAGCTGAGCTTGAATGAATGAACCTGTGGAAGACAAATGAGCTGAGCTTCTGT 3085  
Qy 2810 GCGAGTATATGAAGAACATCTTGTGATCTCAAGACCTGAGATTGTTAAAGTGTATC 2869  
Db 3086 GCGAGTATATGAAGAACATCTTGTGATCTCAAGACCTGAGATTGTTAAAGTGTATC 3145  
Qy 2870 TCACCGCGCGTGTGAGAGTCTGTCCTGTGTGATCTCAAGAGCAGACACCTGAAGA 2929  
Db 3146 TCACCGCGCGTGTGAGAGTCTGTCCTGTGTGATCTCAAGAGCAGACACCTGAAGA 3205



Db 1466 AGAGGTGCCCCCTTCAACCAACGCTCACAGGCTCAGCCCTTTGTGTTCAATC 1525  
Qy 1250 AGTCACCCCTCGAGGCGGTGCGGCGCTGCTCAATCTGAGGAAAGATGTCGA 1309  
Db 1526 AGCTACCCCTCGAGGCGGTGCGGCGCTGCTCAATCTGAGGAAAGATGTCGA 1585  
Qy 1310 AGCGCTTCTGCGATGAGCTGTGAGGAGATGTGAAATAGAAAGTCAAGTGTGATG 1369  
Db 1586 AGCGCTTCTGCGATGAGCTGTGAGGAGATGTGAAATAGAAAGTCAAGTGTGATG 1645  
Qy 1370 AGGACTCATGTTCAAGGATCTGGGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1429  
Db 1646 AGGACTCATGTTCAAGGATCTGGGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1705  
Qy 1430 TCCCTTCCAGAGAGGAGCTGTGAGAGTACTACACCTTCTCCAGCTCAGTCCAGG 1489  
Db 1706 TCCCTTCCAGAGAGGAGCTGTGAGAGTACTACACCTTCTCCAGCTCAGTCCAGG 1765  
Qy 1490 ACTTCTGCGCGCTTGTACTACGTGTAGAGGCGCTGAAATCGAGCAGCTCTGCGC 1549  
Db 1766 ACTTCTGCGCGCTTGTACTACGTGTAGAGGCGCTGAAATCGAGCAGCTCTGCGC 1825  
Qy 1550 CTCTGTAGTTGAGAAACAAAGAGCTCATGAGCTTAAACAGGCGCTTCCATATCC 1609  
Db 1826 CTCTGTAGTTGAGAAACAAAGAGCTCATGAGCTTAAACAGGCGCTTCCATATCC 1885  
Qy 1610 ACTGCTTTGAGATGAAGGCTTCTGTTTGGCTGTGAGGAGAAACGTAGAGAGCCAC 1669  
Db 1886 ACTGCTTTGAGATGAAGGCTTCTGTTTGGCTGTGAGGAGAAACGTAGAGAGCCAC 1945  
Qy 1670 TGAAGGTCCTGTGGGCTGTCCCGTTCCTGTGGGGGTGAGAGAGCTTCTGACTGG 1729  
Db 1946 TGAAGGTCCTGTGGGCTGTCCCGTTCCTGTGGGGGTGAGAGAGCTTCTGACTGG 2005  
Qy 1730 TCTCTGTGTGGGTGAGAGCTTAATGCCACACCCAGAGACACCTTGAAGCTTCC 1789  
Db 2006 TCTCTGTGTGGGTGAGAGCTTAATGCCACACCCAGAGACACCTTGAAGCTTCC 2065  
Qy 1790 ACTGCTTTGAGAGCTCAAGCAAGAAAGTTGTGCGTGGCATTAAGAGCTTCCAG 1849  
Db 2066 ACTGCTTTGAGAGCTCAAGCAAGAAAGTTGTGCGTGGCATTAAGAGCTTCCAG 2125  
Qy 1850 AAGTGTGCTTCGATTAACAGAACCTGACTTATAGATCTTCTTCTGCTCCAGC 1909  
Db 2126 AAGTGTGCTTCGATTAACAGAACCTGACTTATAGATCTTCTTCTGCTCCAGC 2185  
Qy 1910 ACTGTCCGTATTTGCGAAATTCGGGTGATGTCAAGAGATCTTCCAAAGATGAGT 1969  
Db 2186 ACTGTCCGTATTTGCGAAATTCGGGTGATGTCAAGAGATCTTCCAAAGATGAGT 2245  
Qy 1970 CCGCTGAGAGCATGCTGTGTGCTCTATGAGATCGGGATTAAGACCTCATTTAGAGAGC 2029  
Db 2246 CCGCTGAGAGCATGCTGTGTGCTCTATGAGATCGGGATTAAGACCTCATTTAGAGAGC 2305  
Qy 2030 AGTGGGAGAGATTTTGTGCTCAATGTGGCAACCAACCACTGTGGAGAGCTGAGCTGG 2089  
Db 2306 AGTGGGAGAGATTTTGTGCTCAATGTGGCAACCAACCACTGTGGAGAGCTGAGCTGG 2365  
Qy 2090 GCAGCAGATCTCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAGGACATCCA 2149  
Db 2366 GCAGCAGATCTCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAGGACATCCA 2425  
Qy 2150 CCTGTAAATATACAGACCTGTGATTTAGAAATGCAAGATTAACCTTGTGTGACAGAC 2209  
Db 2426 CCTGTAAATATACAGACCTGTGATTTAGAAATGCAAGATTAACCTTGTGTGACAGAC 2485  
Qy 2210 TCTGAGATGCTGATGGCAACCGTAACTTAAGATCCCTCACTTTGGAGGCAACCCAGC 2269  
Db 2486 TCTGAGATGCTGATGGCAACCGTAACTTAAGATCCCTCACTTTGGAGGCAACCCAGC 2545  
Qy 2270 TGAAGGAAGAGATGTAAAGATGGCTGTGAGAGCTTAAACACCCAAATGTTTGTGG 2329

Db 2546 TGAAGGAAGAGATGTAAAGATGGCTGTGAGAGCTTAAACACCCAAATGTTTGTGG 2605  
Qy 2330 AGCTTTGAGAGGCTGATTTGTGATGATGACCCATGCTGTAACTTAAGATCTCCAAA 2389  
Db 2606 AGCTTTGAGAGGCTGATTTGTGATGATGACCCATGCTGTAACTTAAGATCTCCAAA 2665  
Qy 2390 TCTTTAGACCTCCCGAGCTGAAATCTGTGAGCTTGTGGAGAAACAGGTGACAGACC 2449  
Db 2666 TCTTTAGACCTCCCGAGCTGAAATCTGTGAGCTTGTGGAGAAACAGGTGACAGACC 2725  
Qy 2450 AGGAGATTAATGCTCTGAGATGATGCTGAGAGATCTCCAGTGGGCTGTGAGAGAGCTGA 2509  
Db 2726 AGGAGATTAATGCTCTGAGATGATGCTGAGAGATCTCCAGTGGGCTGTGAGAGAGCTGA 2785  
Qy 2510 TACTGAGAGACTGTGAGATCAACAGCAGGATCCAGAGTCTGAGCTCAGCCTCGTCA 2569  
Db 2786 TACTGAGAGACTGTGAGATCAACAGCAGGATCCAGAGTCTGAGCTCAGCCTCGTCA 2845  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTATCCAAACAACGCTGGGAAACGAGGTGTA 2629  
Db 2846 GCAACCGGAGCTTGAACAACCTGTGCTATCCAAACAACGCTGGGAAACGAGGTGTA 2905  
Qy 2630 ATCTACTGTGCTCATCCATGAGGCTTCCCACTGTACTCTGACAGAGCTGATGCTGATC 2689  
Db 2906 ATCTACTGTGCTCATCCATGAGGCTTCCCACTGTACTCTGACAGAGCTGATGCTGATC 2965  
Qy 2690 AGTGCACCTGGAACAGGCTGTGAGTGTCTGTGACTTGTGACCTTATGGGTAACTCAT 2749  
Db 2966 AGTGCACCTGGAACAGGCTGTGAGTGTCTGTGACTTGTGACCTTATGGGTAACTCAT 3025  
Qy 2750 GGTGACGACCTGAGCTTATGACATGAACCTGTGAAAGACATGCGGTGAAGCTTCTGT 2809  
Db 3026 GGTGACGACCTGAGCTTATGACATGAACCTGTGAAAGACATGCGGTGAAGCTTCTGT 3085  
Qy 2810 GCGAGGTCAATGAGAAACCATTTGTGATCTCCAGAGACTGTGAGATTTGTAAGTGTATC 2869  
Db 3086 GCGAGGTCAATGAGAAACCATTTGTGATCTCCAGAGACTGTGAGATTTGTAAGTGTATC 3145  
Qy 2870 TCACCGCGCGGTGCTGTGAGAGTCTGTCTGTGATCTCGAGGAGAGACACTGAAGA 2929  
Db 3146 TCACCGCGCGGTGCTGTGAGAGTCTGTCTGTGATCTCGAGGAGAGACACTGAAGA 3205  
Qy 2930 GCTGTGATCTTCAACGAAACATGCTGTGGGTGAACGCTGGGCTGTGCGAGGAGAC 2989  
Db 3206 GCTGTGATCTTCAACGAAACATGCTGTGGGTGAACGCTGGGCTGTGCGAGGAGAC 3265  
Qy 2990 TGAAGCAAAAGAACAGTGTCTGACAGAGACTCGGCTTGAAGGCAATGTGACTGACTCTG 3049  
Db 3366 TGAAGCAAAAGAACAGTGTCTGACAGAGACTCGGCTTGAAGGCAATGTGACTGACTCTG 3325  
Qy 3050 ATTGCTGTGAGGCACTGCTCCCTTGGCCCTTCTGCAACCCGCACTGACAGCTTAACC 3109  
Db 3326 ATTGCTGTGAGGCACTGCTCCCTTGGCCCTTCTGCAACCCGCACTGACAGCTTAACC 3385  
Qy 3110 TGTGTGAGATTAATCTTCAAGTCCCAAGAAATGATGAAGCTGTGTGCGCTTTGCGCTGTC 3169  
Db 3386 TGTGTGAGATTAATCTTCAAGTCCCAAGAAATGATGAAGCTGTGTGCGCTTTGCGCTGTC 3445  
Qy 3170 CCACGTCTAATTAACAGATTAATTTGGCTGTGAAATGACAGTACTTGTGCAATTAAGGA 3229  
Db 3446 CCACGTCTAATTAACAGATTAATTTGGCTGTGAAATGACAGTACTTGTGCAATTAAGGA 3505  
Qy 3230 AGCTGTGAGAGAGTCAAGCTCAAGCCCGAGCTGTAATTTGACAGGTGATTTGGGATTT 3289  
Db 3506 AGCTGTGAGAGAGTCAAGCTCAAGCCCGAGCTGTAATTTGACAGGTGATTTGGGATTT 3565  
Qy 3290 CTTTGAATGAAGATGACCG 3308  
Db 3566 CTTTGAATGAAGATGACCG 3584

RESULT 12  
ABK48609

ID ABK48609 standard; cDNA; 1157 BP.  
XX  
AC ABK48609;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Human MATER cDNA fragment #1.  
XX  
KW Human; gene; 88; contraceptive; antifertility; MATER;  
KW Maternal antigen that embryos require; MATER null phenotype; oocyte;  
KW early embryonic survival; premature ovarian failure; POF;  
KW autoimmune infertility; chromosome 19; gene therapy; fertility.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1155  
FT /tag= a  
FT /product= "Human MATER"  
FT /partial  
FT /note= "No start or stop codon shown"  
XX  
PN WO200232955-A1.  
XX  
PD 25-APR-2002.  
XX  
PP 04-APR-2001; 2001WO-US010981.  
XX  
PR 18-OCT-2000; 2000US-0241510P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nelson LM, Tong Z;  
XX  
DR MPI; 2002-454595/48.  
XX  
DR P-PsDB; AAU79523.  
XX  
PT New isolated human Maternal Antigen That Embryos Require protein and  
PT polynucleotide, useful in diagnosing or treating fertility or reduced  
PT fertility, or as a contraceptive.  
XX  
PS Claim 11; Page 68-69; 93pp; English.  
XX  
CC The invention discloses an isolated human MATER (Maternal Antigen That  
CC Embryos Require) protein, which can complement a MATER null phenotype in  
CC which zygotes arising from the oocyte do not progress beyond the two-cell  
CC stage. MATER is required for early embryonic survival and abnormal levels  
CC of the protein can lead to premature ovarian failure (POF) and can be  
CC caused by under or over expression of MATER or an autoimmune response to  
CC MATER. MATER is a single-copy Maternal effect gene found on chromosome  
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in  
CC diagnosing or treating fertility and reduced fertility. In particular,  
CC the MATER protein is useful as a contraceptive agent, or for influencing  
CC (either inhibiting or enhancing) fertility and can be used to detect a  
CC predisposition to infertility or reduced fertility, or for presymptomatic  
CC screening of an individual for infertility/reduced fertility. The protein  
CC and polynucleotide are also useful for detecting an excess or deficiency,  
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a  
CC human or a mouse) or for screening for a compound useful in influencing  
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA  
CC fragment #1, which was isolated from a human cDNA library  
XX  
SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;  
Query Match 27.3%; Score 1102; DB 6; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 757 GCTCCGATGACGAGATCATGTCCTCCGACGAGAGGCTGTTTATCATATTGACGGTTTC 816  
DB 4 GCTCCGATGACGAGATCATGTCCTCCGACGAGAGGCTGTTTATCATATTGACGGTTTC 63  
QY 817 GATGACCTGGGCTGTCTCTCAATGACAGAAAGCTCTGCAAAAGACTGGGCTGAGAG 876

DB 64 GATGACCTGGGCTGTCTCTCAATGACAGAAAGCTCTGCAAAAGACTGGGCTGAGAG 123  
QY 877 CAGCTCCGTTACACCTCATACGACAGTCTGTGAGAGAAAGTCTGCTCCGTAGTCTTC 936  
DB 124 CAGCTCCGTTACACCTCATACGACAGTCTGTGAGAGAAAGTCTGCTCCGTAGTCTTC 183  
QY 937 CTGATCTGACACCTGACAGACCTGGGACACAGAAAGCTCAAGTCAAGAGTCTGTCTCC 996  
DB 184 CTGATCTGACACCTGACAGACCTGGGACACAGAAAGCTCAAGTCAAGAGTCTGTCTCC 243  
QY 997 CATTACCTGTTAGTTAGAGAAATCTCCGGGAGAAAGAAATCACTTGTCTTAGAGCC 1056  
DB 244 CATTACCTGTTAGTTAGAGAAATCTCCGGGAGAAAGAAATCACTTGTCTTAGAGCC 303  
QY 1057 GGGATTTGGTGAACATCAGAAAGACAGAAAGGTTCCGTGCGATCATGAACACCGTAGCTG 1116  
DB 304 GGGATTTGGTGAACATCAGAAAGACAGAAAGGTTCCGTGCGATCATGAACACCGTAGCTG 363  
QY 1117 CTGACACAGTGCAGAGTCCCGGCGGTGAGTCTCTCATCTGTGAGGCTCTGAGCTGCAG 1176  
DB 364 CTGACACAGTGCAGAGTCCCGGCGGTGAGTCTCTCATCTGTGAGGCTCTGAGCTGCAG 423  
QY 1177 GACGTGTGGGGAGAGGCGTCCGCTTCAACCAAGGCTCAAGGCTGACGCGCTG 1236  
DB 424 GACGTGTGGGGAGAGGCGTCCGCTTCAACCAAGGCTCAAGGCTGACGCGCTG 483  
QY 1237 TTTGTGTTTCATCAGCTCAACCCCTTGAGGCGTGTCCGGCGCTGTCTCATCTGAGGAA 1296  
DB 484 TTTGTGTTTCATCAGCTCAACCCCTTGAGGCGTGTCCGGCGCTGTCTCATCTGAGGAA 543  
QY 1297 AGAGTTGCTCGAAGCGCTTCTGCGGTATGCTGTGAGAGGAGTGTGAATGGAATGCA 1356  
DB 544 AGAGTTGCTCGAAGCGCTTCTGCGGTATGCTGTGAGAGGAGTGTGAATGGAATGCA 603  
QY 1357 GGTGTTGAGTGAAGACCTCATGTTGTAAGGACCTGGGGAGTCAAGCTCCGCTCTG 1416  
DB 604 GGTGTTGAGTGAAGACCTCATGTTGTAAGGACCTGGGGAGTCAAGCTCCGCTCTG 663  
QY 1417 TTTGACATGAACATCTTCTCCAGACAGCACTGTGAGAGTACTACCTTCTTCAC 1476  
DB 664 TTTGACATGAACATCTTCTCCAGACAGCACTGTGAGAGTACTACCTTCTTCAC 723  
QY 1477 CTGACTCTCAGAGCTTCTGTGCGGCTTGTACTAGTGTGAGAGGCTGAGAAATCGAG 1536  
DB 724 CTGACTCTCAGAGCTTCTGTGCGGCTTGTACTAGTGTGAGAGGCTGAGAAATCGAG 783  
QY 1537 CCAAGTCTGCGGCTCTGTGAGTGAAGAAAGAGAGTCCATGAGACTTAAACAGGCA 1596  
DB 784 CCAAGTCTGCGGCTCTGTGAGTGAAGAAAGAGAGTCCATGAGACTTAAACAGGCA 843  
QY 1597 GGCCTTCATATCACTGCTTGTGATGAAGAGTCTTGTGAGGCTGTGAGGAGAAAG 1656  
DB 844 GGCCTTCATATCACTGCTTGTGATGAAGAGTCTTGTGAGGCTGTGAGGAGAAAG 903  
QY 1657 GTAAGAGAGCACTGAGAGTCTGCTGAGGCTGTGAGGCTGTGAGGAGAAAG 1716  
DB 904 GTAAGAGAGCACTGAGAGTCTGCTGAGGCTGTGAGGCTGTGAGGAGAAAG 963  
QY 1717 CTTCTGCACTGGGCTCTCTGTGGGTGAAGAGCTTAATGCAACACCCAGAGACACC 1776  
DB 964 CTTCTGCACTGGGCTCTCTGTGGGTGAAGAGCTTAATGCAACACCCAGAGACACC 1023  
QY 1777 CTGAGAGGCTTCACTGCTTCTTGAAGCTCAAGAGAGTGTGCTGCTGGCATTTA 1836  
DB 1024 CTGAGAGGCTTCACTGCTTCTTGAAGCTCAAGAGAGTGTGCTGCTGGCATTTA 1083  
QY 1837 AACAGCTTCAAGAGAGTGTGCTTCAATTAACAGAACTGAGCTTGATGATCTTCC 1896  
DB 1084 AACAGCTTCAAGAGAGTGTGCTTCAATTAACAGAACTGAGCTTGATGATCTTCC 1143  
QY 1897 TTTGCTCCGACG 1909

Dh	1144	TTCGCGCTCCAGC	1156
<b>RESULT 13</b>			
AC	AAD48999	AAD48999 standard; cDNA; 1157 BP.	
XX	AAD48999;		
DT	07-MAR-2003	(first entry)	
XX	Human MATER cDNA fragment 1.		
KW	Human; MATer protein; infertility; fertility; contraceptive agent;		
KW	gene therapy; gene; ss.		
XX	Homo sapiens.		
XX			
Key	Location/Qualifiers		
CDS	1..1155		
FT	/*tag= a		
FT	/product= "Human MATER protein fragment 1"		
FT	/transl_except= (pos:112..114, aa:Thr)		
FT	/transl_except= (pos:589..591, aa:Thr)		
FT	/transl_except= (pos:865..867, aa:Thr)		
FT	/transl_except= (pos:973..975, aa:Thr)		
FT	/transl_except= (pos:1102..1104, aa:Thr)		
FT	/note= "No start and stop codon"		
PT	/partial		
PX	WO200281492-A1.		
XX			
PD	17-OCT-2002.		
XX			
PF	29-MAR-2002; 2002WO-US009776.		
XX			
PR	04-APR-2001; 2001WO-US010981.		
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
PA	Nelson LM, Tong Z;		
P1	WP1; 2003-058494/05.		
DR	P-P5DB; AAB31746.		
XX			
XX	New isolated variant MATER proteins and nucleic acids, useful for		
PT	diagnosing, prognosing and treating infertility and reduced fertility,		
PT	and as contraceptive agents.		
XX			
XX	Example 2; Page 79-80; 110pp; English.		
PS			
FS	The present invention relates to novel MATER proteins and polynucleotides		
CC	encoding such proteins. The MATER proteins are essential to fertility.		
CC	Sequences of the invention are useful for diagnosing, prognosing and		
CC	treating infertility, reduced fertility and as contraceptive agents. They		
CC	are also useful in gene therapy. The method is useful for detecting a		
CC	predisposition to or pre-symptomatic screening of an individual for		
CC	infertility or reduced fertility. The present sequence is human MATER		
CC	cDNA fragment 1		
SQ			
Sequence	1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;		
Query Match	27.3%; Score 1102; DB 8; Length 1157;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1152;	Conservative 0; Mismatches 1; Indels 0; Gaps 0		
OY	757 GCTCGGTGACGGAGATCATGTCCGCACAAAGGCTGTGTTTCATTGACGGTTC	816	
Db	4 GCTCGGTGACGGAGATCATGTCCGCACAAAGGCTGTGTTTCATTGACGGTTC	63	
OY	817 GATGACCTGGGCTCTGTCTCAACATGACACAAAGCTCTGAAAAGATGGGCTGAAGA	876	
Db	64 GATGACCTGGGCTCTGTCTCAACATGACACAAAGCTCTGAAAAGATGGGCTGAAGA	123	

QY	877	TAAGCTCCGTTTACCCCTCATPACGAGCTGTGTGAGGAAAGGCTCGAGTCCCGAGTCCCTC	936
Db	124	CAGCTCCGTTACCCCTCATPACGAGCTGTGTGAGGAAAGGCTGTCTCCCTGAGTCTTTC	183
QY	937	CTGATCCGTCAACCGTCAGAGAGCTGGGACACAGAGAAGCTCAAGTCAGAGGTGTGTCTCC	996
Db	184	CTGATCCGTCAACCGTCAGAGAGCTGGGACACAGAGAAGCTCAAGTCAGAGGTGTGTCTCC	243
QY	997	CGTTACCTGTTAGTTAGAGAAATCTCCGGGGAACAAAGATCCACTTGCTCTTGGAGCGC	1056
Db	244	CGTTACCTGTTAGTTAGAGAAATCTCCGGGGAACAAAGATCCACTTGCTCTTGGAGCGC	303
QY	1057	GGGATTTGGTAGACATCAGAAAGACACAGAGGTTGGTGCGATCATGAACAAACCGTAGACTG	1116
Db	304	GGGATTTGGTAGACATCAGAAAGACACAGAGGTTGGTGCGATCATGAACAAACCGTAGACTG	363
QY	1117	CTCCACCAAGTCCACAGTGCACCGCGTGGGCTCTCTCACTTGCCTGGCCCTGCACCTGCAG	1176
Db	364	CTCCACCAAGTCCACAGTGCACCGCGTGGGCTCTCTCACTTGCCTGGCCCTGCACCTGCAG	423
QY	1177	GACGTGTGGGGAGAGAGTGCCTTCCTTCAACCAAAAGCTCACAGGCTGCACGCGCT	1236
Db	424	GACGTGTGGGGAGAGAGTGCCTTCCTTCAACCAAAAGCTCACAGGCTGCACGCGCT	483
QY	1237	TTTGTGTTCATCAGCTCAACCCCTCGAGAGGTGGTCCGGCGCTGTCTCAATCTGAGGAA	1296
Db	484	TTTGTGTTCATCAGCTCAACCCCTCGAGAGGTGGTCCGGCGCTGTCTCAATCTGAGGAA	543
QY	1297	AGAGTGTTCGGAAGCGCTTCTGCAGATGAGCTGTGAGGGAGTGGAAATGAGAAATCA	1356
Db	544	AGAGTGTTCGGAAGCGCTTCTGCAGATGAGCTGTGAGGGAGTGGAAATGAGAAATCA	603
QY	1357	GTTGTTGATGTGTGACGAACTTCATGATTCAAAGAACTCGGAGTGTGACTTCGGTCTTG	1416
Db	604	GTTGTTGATGTGTGACGAACTTCATGATTCAAAGAACTCGGAGTGTGACTTCGGTCTTG	663
QY	1417	TTTCAATGAACATCTCTTCTCCAGACACCACTGTGAGAGTACTACACTTCTTCCAC	1478
Db	664	TTTCAATGAACATCTCTTCTCCAGACACCACTGTGAGAGTACTACACTTCTTCCAC	723
QY	1477	CTCAGTCCCAAGGACTCTGTGGCGCGCTGTGACTACGTTGTAGAGGGCTGGAATCGAG	1536
Db	724	CTCAGTCCCAAGGACTCTGTGGCGCGCTGTGACTACGTTGTAGAGGGCTGGAATCGAG	783
QY	1537	CCAGCTCTTGCCTCTGTACGTTGAGAAAGACAAAGAGGTCATGAGACTTTAAACAGCA	1596
Db	784	CCAGCTCTTGCCTCTGTGTGAGTGAAGAAAGCAAGAGGTCATGAGACTTTAAACAGCA	843
QY	1597	GCGTTCCAAATCACTGCTTGTGATGAAAGCGTTTCTTGTGGCTCTGTAGCGAAGAC	1656
Db	844	GCGTTCCAAATCACTGCTTGTGATGAAAGCGTTTCTTGTGGCTCTGTAGCGAAGAC	903
QY	1657	GTAAGAGAGCCACGTGAGAGTCTGTGGGCTGTCCCTTCCCTGGGGGTGAACAGAG	1716
Db	904	GTAAGAGAGCCACGTGAGAGTCTGTGGGCTGTCCCTTCCCTGGGGGTGAACAGAG	963
QY	1717	CTTTCGACACTGGGCTCTGTGTTGGGTACAGAGCGCTAATGCAACCAACCCAGAGACAC	1776
Db	964	CTTTCGACACTGGGCTCTGTGTTGGGTACAGAGCGCTAATGCAACCAACCCAGAGACAC	1023
QY	1777	CTGAGAGCGCTTCACTGTCTTTTGGAGCTTCAGACAAAGATTTGTTCGTTGGCATTA	1836
Db	1024	CTGAGAGCGCTTCACTGTCTTTTGGAGCTTCAGACAAAGATTTGTTCGTTGGCATTA	1083
QY	1837	AAACAGCTTCCAAAGATGTGGCTTCCGATTAACAGAACTGTGACTTGATAGCATTTCC	1896
Db	1084	AAACAGCTTCCAAAGATGTGGCTTCCGATTAACAGAACTGTGACTTGATAGCATTTCC	1143
QY	1897	TTTCGGCTCCAGC 1909	
Db	1144	TTTCGGCTCCAGC 1156	



RESULT 14  
 ABK48610  
 ID ABK48610 standard; cDNA; 1075 BP.  
 XX  
 AC ABK48610;  
 XX  
 DT 13-AUG-2002 (first entry)  
 XX  
 DE Human MATER cDNA fragment #2.  
 XX  
 KW Human; gene; ss; contraceptive; antiinfertility; MATER;  
 KW maternal antigen that embryos require; MATER null phenotype; oocyte;  
 KW early embryonic survival; premature ovarian failure; POF;  
 KW autoimmune infertility; chromosome 19; gene therapy; fertility.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 2..778  
 FT /tag= a  
 FT /product= "Human MATER"  
 FT /partial  
 FT /note= "No start codon shown"  
 XX  
 PN W0200232955-A1.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 04-APR-2001; 2001WO-US010981.  
 XX  
 PR 18-OCT-2000; 2000US-0241510P.  
 XX  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Nelson LM, Tong Z;  
 XX  
 DR WPI; 2002-454595/48.  
 XX  
 DR P-PSDB; AAU79524.  
 XX  
 PT New isolated human Maternal Antigen That Embryos Require protein and  
 PT polynucleotide, useful in diagnosing or treating fertility or reduced  
 PT fertility, or as a contraceptive.  
 XX  
 PS Claim 11; Page 70-71; 93pp; English.  
 XX  
 CC The invention discloses an isolated human MATER (Maternal Antigen That  
 CC Embryos Require) protein, which can complement a MATER null phenotype in  
 CC which zygotes arising from the oocyte do not progress beyond the two-cell  
 CC stage. MATER is required for early embryonic survival and abnormal levels  
 CC of the protein can lead to premature ovarian failure (POF) and can be  
 CC caused by under or over expression of MATER or an autoimmune response to  
 CC MATER. MATER is a single-copy maternal effect gene found on chromosome  
 CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in  
 CC diagnosing or treating fertility and reduced fertility. In particular,  
 CC the MATER protein is useful as a contraceptive agent, or for influencing  
 CC (either inhibiting or enhancing) fertility and can be used to detect a  
 CC predisposition to infertility or reduced fertility, or for presymptomatic  
 CC screening of an individual for infertility/reduced fertility. The protein  
 CC and polynucleotide are also useful for detecting an excess or deficiency,  
 CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a  
 CC human or a mouse) or for screening for a compound useful in influencing  
 CC MATER-mediated fertility. The sequence presented is the human MATER cDNA  
 CC fragment #2, which was isolated from a human cDNA library  
 XX  
 SO Sequence 1075 BP; 278 A; 251 C; 281 G; 265 T; 0 U; 0 Other;  
 XX  
 Query Match 16.3%; Score 657; DB 6; Length 1075;  
 Best Local Similarity 99.7%; Pred. No. 1,7e-297;  
 Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2550 TCTGGCCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTCCTATCCAAACAG 2609  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 TCTGGCCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTCCTATCCAAACAG 60  
 QY 2610 CCTGGGGAACGAAGGTGTAATCTACTGTGTGATCCATGAGGCTTCCCATGTAGTCT 2669  
 DB 61 CTTGGGGAAACGAAGGTGTAATCTACTGTGTGATCCATGAGGCTTCCCATGTAGTCT 120  
 QY 2670 GCAGAGGCTGATGCTGATCAGTGCACCTGGAACCGCTGGCTGTGTTTCTTGCACT 2729  
 DB 121 GCAGAGGCTGATGCTGATCAGTGCACCTGGAACCGCTGGCTGTGTTTCTTGCACT 180  
 QY 2730 TGCCTTATGGGTAACTCATGAGCTGACGACCTGAGCTTATGCAATGAACTGTGGAAGA 2789  
 DB 181 TGCCTTATGGGTAACTCATGAGCTGACGACCTGAGCTTATGCAATGAACTGTGGAAGA 240  
 QY 2790 CAATGGCCTGAAAGCTTCTGTGCGAGTGCATAGAGAACCACTTGTCATCTCGAGACT 2849  
 DB 241 CAATGGCCTGAAAGCTTCTGTGCGAGTGCATAGAGAACCACTTGTCATCTCGAGACT 300  
 QY 2850 GAAGTTGGTAAAGTGCATCTCACCGCGCTGTGAGAGTCTGTGCTGTGATCTC 2909  
 DB 301 GAAGTTGGTAAAGTGCATCTCACCGCGCTGTGAGAGTCTGTGCTGTGATCTC 360  
 QY 2910 GAGAGCAGACACCTGAAAGAGCTGATCTCACGACCAATGCTTGGGTGACGGTGGGT 2969  
 DB 361 GAGAGCAGACACCTGAAAGAGCTGATCTCACGACCAATGCTTGGGTGACGGTGGGT 420  
 QY 2970 TGTCTGCGCTGTGCGAGGAGCTGAAAGCAAAAGATTTCTGACGAACTCGGGTTGAA 3029  
 DB 421 TGTCTGCGCTGTGCGAGGAGCTGAAAGCAAAAGATTTCTGACGAACTCGGGTTGAA 480  
 QY 3030 GGCAATGGAGACCTTCTGATGCTGTGAGGACCTCTTGGCCCTTCTCGCAACG 3089  
 DB 481 GGCAATGGAGACCTTCTGATGCTGTGAGGACCTCTTGGCCCTTCTCGCAACG 540  
 QY 3090 GCATCTGACCAAGTCAAACTGTGACAGATTAATCTTACGCTCCAAAGGAATGATGAGCT 3149  
 DB 541 GCATCTGACCAAGTCAAACTGTGACAGATTAATCTTACGCTCCAAAGGAATGATGAGCT 600  
 QY 3150 GTGTTGGGCTTGTGCTGTCCAGGTCTTAATCTTACGATTAATGGCTGTGAAATGCA 3209  
 DB 601 GTGTTGGGCTTGTGCTGTCCAGGTCTTAATCTTACGATTAATGGCTGTGAAATGCA 660  
 QY 3210 GTACCTGTGCAAAATTAAGAGCTGTGAGAGAGTGACGTAATCAAGCCCGAGTGT 3269  
 DB 661 GTACCTGTGCAAAATTAAGAGCTGTGAGAGAGTGACGTAATCAAGCCCGAGTGT 720  
 QY 3270 AATTGACGTAGTGGCATTTCTTTGATGATGATGATGACCG 3308  
 DB 721 AATTGACGTAGTGGCATTTCTTTGATGATGATGATGACCG 759  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 15  
 AAD49000  
 ID AAD49000 standard; cDNA; 1075 BP.  
 XX  
 AC AAD49000;  
 XX  
 DT 07-MAR-2003 (first entry)  
 XX  
 DE Human MATER cDNA fragment 2.  
 XX  
 KW Human; MATER protein; infertility; fertility; contraceptive agent;  
 KW gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..778  
 FT /tag= a  
 FT /product= "Human MATER protein fragment 2"  
 FT /transl\_except= (pos:200..202, aa:Thr)  
 FT /transl\_except= (pos:650..652, aa:Thr)  
 FT /transl\_except= (pos:656..658, aa:Thr)

```
FT      /transl_except= (pos:734, .736, aa:Thr)
FT      /transl_except= (pos:764, .766, aa:Thr)
FT      /transl_except= (pos:767, .769, aa:Thr)
FT      /note="No start codon"
FT      /partial
XX      WO200281492-A1.
XX      17-OCT-2002.
XX      29-MAR-2002; 2002WO-US009776.
XX      04-APR-2001; 2001WO-US010981.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      Nelson LM, Tong Z;
XX      WPI, 2003-058494/05.
XX      P-PSDB; AAE31747.
XX      New isolated variant MATER proteins and nucleic acids, useful for
XX      diagnosing, prognosing and treating infertility and reduced fertility,
XX      and as contraceptive agents.
XX      Example 2; Page 81-82; 110pp; English.
XX      The present invention relates to novel MATER proteins and polynucleotides
XX      encoding such proteins. The MATER proteins are essential to fertility.
XX      Sequences of the invention are useful for diagnosing, prognosing and
XX      treating infertility, reduced fertility and as contraceptive agents. They
XX      are also useful in gene therapy. The method is useful for detecting a
XX      predisposition to or pre-symptomatic screening of an individual for
XX      infertility or reduced fertility. The present sequence is human MATER
XX      CDNA fragment 2
XX      SQ      Sequence 1075 BP; 278 A; 251 C; 281 G; 265 T; 0 U; 0 Other;
XX
XX      Query Match      16.3%; Score 657; DB 8; Length 1075;
XX      Best Local Similarity 99.7%; Pred. No. 1.7e-297;
XX      Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY      2550 TCTGGCCTCAGCCCTGCTGAGCAACCGAGCTTGACACACCTGTGCTATCCAAACACAG 2609
DB      1 TCTGGCCTCAGCCCTGCTGAGCAACCGAGCTTGACACACCTGTGCTATCCAAACACAG 60
QY      2610 CTTGGGGAACGAGGTTAAATCTACTGTGTGATCCATGAGGCTTCCCACTGATGCT 2669
DB      61 CTTGGGGAACGAGGTTAAATCTACTGTGTGATCCATGAGGCTTCCCACTGATGCT 120
QY      2670 GCAGAGCTGATGCTGATCAAGTGCACCTGGACACGCGCTGTGTGTTTCTTGCACT 2729
DB      121 GCAGAGCTGATGCTGATCAAGTGCACCTGGACACGCGCTGTGTGTTTCTTGCACT 180
QY      2730 TGGCTTATGGGTAATCTATGCTGACGACCTGAGCCTTAAGCATGAAACCTGTGGAAGA 2789
DB      181 TGGCTTATGGGTAATCTATGCTGACGACCTGAGCCTTAAGCATGAAACCTGTGGAAGA 240
QY      2790 CAATGGCGTGAAGCTTCTGTGAGGTATGAGAGAACATCTTGTCACTCCAGAACT 2849
DB      241 CAATGGCGTGAAGCTTCTGTGAGGTATGAGAGAACATCTTGTCACTCCAGAACT 300
QY      2850 GGAAGTTGTTAAAGTCTCATCTCACCGCGCGCTGTGAGAGTCTGTCTGTGATCTC 2909
DB      301 GGAAGTTGTTAAAGTCTCATCTCACCGCGCGCTGTGAGAGTCTGTCTGTGATCTC 360
QY      2910 GAGAGCAACAACCTGAAGAGCTGATCTCACGACCAATGCGCTGGGTGACGCTGGGGT 2969
DB      361 GAGAGCAACAACCTGAAGAGCTGATCTCACGACCAATGCGCTGGGTGACGCTGGGGT 420
QY      2970 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGACGTTTCTGACGAGACTCGGGTTGAA 3029
DB      421 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGACGTTTCTGACGAGACTCGGGTTGAA 480
```

```
QY      3030 GGCAATGAGACTGACTTCTGATGCTGTGAGGACACTCTCCTTGGCCCTTCTGCAACCG 3089
DB      481 GGCAATGAGACTGACTTCTGATGCTGTGAGGACACTCTCCTTGGCCCTTCTGCAACCG 540
QY      3090 GCATCTGACCAAGTCTAAACCTGTGTGCAAGATTAATCTTACCTCCAAAGAGATGAAAGCT 3149
DB      541 GCATCTGACCAAGTCTAAACCTGTGTGCAAGATTAATCTTACCTCCAAAGAGATGAAAGCT 600
QY      3150 GTGTTGGGCTTGTGCTGTGCTCCAGTCTTAAGTAAATGAGGCTGTGGAATGAGCA 3209
DB      601 GTGTTGGGCTTGTGCTGTGCTCCAGTCTTAAGTAAATGAGGCTGTGGAATGAGCA 660
QY      3210 GTACCTGTGCAATTAAGAGCTGTGAGAGAGTCAAGTCAAGCTCAAGCCCGAGTCTG 3269
DB      661 GTACCTGTGCAATTAAGAGCTGTGAGAGAGTCAAGTCAAGCTCAAGCCCGAGTCTG 720
QY      3270 AATTGACGGTATGTTGGCATCTTTTGAATGAAGATGACCG 3308
DB      721 AATTGACGGTATGTTGGCATCTTTTGAATGAAGATGACCG 759
```

Search completed: July 19, 2005, 03:00:34  
Job time : 2018 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2005, 01:48:06 ; Search time 633 Seconds

(without alignments)  
10430.295 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atggaagagagacaatcgctc.....tcgaggcgctggtccttaa 4035

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 81813359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCDS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	0.5	601	US-09-949-016-91585	Sequence 91585, A
C 2	20	0.5	601	US-09-949-016-91586	Sequence 91586, A
C 3	20	0.5	6042	US-08-261-822A-1	Sequence 1, Appl
C 4	20	0.5	6042	PCT-US95-07744A-1	Sequence 1, Appl
C 5	20	0.5	6172	US-08-819-288-1	Sequence 1, Appl
C 6	20	0.5	6172	US-10-144-198-11	Sequence 11, Appl
C 7	20	0.5	8733	US-09-949-016-13371	Sequence 14371, A
C 8	20	0.5	57837	US-09-949-016-17601	Sequence 17601, A
C 9	20	0.5	57839	US-09-949-016-17601	Sequence 17601, A
C 10	19	0.5	400	US-09-621-976-2835	Sequence 2835, Ap
C 11	19	0.5	601	US-09-949-016-151197	Sequence 151197, A
C 12	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 13	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 14	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 15	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 16	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 17	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 18	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 19	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 20	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 21	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 22	19	0.5	1110	US-09-543-681A-2553	Sequence 2553, Ap
C 23	19	0.5	1341	US-09-513-999C-14927	Sequence 14927, A
C 24	19	0.5	1428	US-09-248-796A-3252	Sequence 3252, Ap
C 25	19	0.5	2848	US-09-023-655-1364	Sequence 1364, Ap
C 26	19	0.5	3180	US-09-248-796A-20	Sequence 20, Appl
C 27	19	0.5	3349	US-09-375-318-36	Sequence 36, Appl

28	19	0.5	5521	US-08-956-171E-408	Sequence 408, App
29	19	0.5	5521	US-08-781-986A-408	Sequence 408, App
30	19	0.5	13198	US-09-949-016-16425	Sequence 16425, A
31	19	0.5	28843	US-09-949-016-11812	Sequence 11812, A
32	19	0.5	28843	US-09-949-016-12712	Sequence 12712, A
33	19	0.5	28843	US-09-949-016-17158	Sequence 17158, A
34	19	0.5	28843	US-09-949-016-17158	Sequence 17159, A
35	19	0.5	48039	US-09-949-016-15990	Sequence 15990, A
36	19	0.5	49315	US-09-949-016-13016	Sequence 13016, A
37	19	0.5	56326	US-09-949-016-16468	Sequence 16468, A
38	19	0.5	57103	US-09-949-016-13445	Sequence 13445, A
39	19	0.5	60417	US-09-949-016-13312	Sequence 13312, A
40	19	0.5	71251	US-09-949-016-15332	Sequence 15332, A
41	19	0.5	209210	US-09-949-016-15094	Sequence 15094, A
42	19	0.5	278866	US-09-949-016-13922	Sequence 13922, A
43	19	0.5	278866	US-09-949-016-13923	Sequence 13923, A
44	19	0.5	278866	US-09-949-016-13924	Sequence 13924, A
45	19	0.5	278866	US-09-949-016-13925	Sequence 13925, A

## ALIGNMENTS

```
RESULT 1
US-09-949-016-91585/c
; Sequence 91585, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91585
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-91585

Query Match          0.5%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2398 ACCCTCCCGAGCGTGAATC 2417
DB 476 ACCTCCCGAGCGTGAATC 457

RESULT 2
US-09-949-016-91586/c
; Sequence 91586, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 91586  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-91586

Query Match 0.5%; Score 20; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2398 ACCTCCCGAGCTGAATC 2417  
DB 290 ACCTCCCGAGCTGAATC 271

## RESULT 3

US-08-261-822A-1/c  
Sequence 1, Application US/08261822A  
Patent No. 5650553

GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph R. et al.  
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene  
TITLE OF INVENTION: and Pathogens  
NUMBER OF SEQUENCES: 82  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553r1s  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/261,822A  
FILING DATE: 17-JUN-1994  
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6042 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-261-822A-1

Query Match 0.5%; Score 20; DB 1; Length 6042;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 966 AGAGAGCTCAAGTCAGAG 985  
DB 2046 AGAGAGCTCAAGTCAGAG 2027

## RESULT 4

PCT-US95-07744A-1/c  
Sequence 1, Application PC/TUS9507744A  
GENERAL INFORMATION:

APPLICANT: Trustees of The University of Pennsylvania  
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene  
TITLE OF INVENTION: and Pathogens  
NUMBER OF SEQUENCES: 82  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07744A  
FILING DATE: 15-JUNE-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/261,822  
FILING DATE: June 17, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6042 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-07744A-1

Query Match 0.5%; Score 20; DB 5; Length 6042;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 966 AGAGAGCTCAAGTCAGAG 985  
DB 2046 AGAGAGCTCAAGTCAGAG 2027

RESULT 5  
US-08-819-288-1/c  
Sequence 1, Application US/08819288  
Patent No. 5955652

GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph  
APPLICANT: Alonso, Jose  
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE  
TITLE OF INVENTION: AND PATHOGENS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652r1s  
STREET: One Liberty Place - 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/819,288

FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: UPN-2949  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6172 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-819-288-1

Query Match 0.5%; Score 20; DB 2; Length 6172;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 966 AGAGAGCTCAAGTCAGAG 985  
Db 1975 AGAGAGCTCAAGTCAGAG 1956

RESULT 6  
US-09-400-348-1/c  
Sequence 1, Application US/09400348  
Patent No. 6355778  
GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph  
APPLICANT: Alonso, Jose  
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE  
TITLE OF INVENTION: AND PATHOGENS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 635577818  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/400,348  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/819,288  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: UPN-2949  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6172 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO

ANTI-SENSE: NO  
US-09-400-348-1

Query Match 0.5%; Score 20; DB 3; Length 6172;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 966 AGAGAGCTCAAGTCAGAG 985  
Db 1975 AGAGAGCTCAAGTCAGAG 1956

RESULT 7  
US-10-144-198-11  
Sequence 11, Application US/10144198  
Patent No. 6833247  
GENERAL INFORMATION:  
APPLICANT: Origene Technologies Inc  
TITLE OF INVENTION: Regulated Prostate Cancer Genes  
FILE REFERENCE: 9U 105 R1  
CURRENT APPLICATION NUMBER: US/10/144,198  
CURRENT FILING DATE: 2002-05-14  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 8733  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (473)..(8566)  
US-10-144-198-11

Query Match 0.5%; Score 20; DB 4; Length 8733;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 ACAGCAGCAGCAGCAGAGA 341  
Db 8459 ACAGCAGCAGCAGCAGAGA 8478

RESULT 8  
US-09-949-016-14371/c  
Sequence 14371, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14371  
LENGTH: 57837  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-14371

Query Match 0.5%; Score 20; DB 4; Length 57837;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2398 ACCTCCCCAGCCTGAATC 2417  
|||||

Db 19969 ACCTCCCCAGCTGAATC 19950

## RESULT 9

US-09-949-016-17601/c  
; Sequence 17601, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17601  
; LENGTH: 57839  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17601

Query Match 0.5%; Score 20; DB 4; Length 57839;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2398 ACCTCCCCAGCTGAATC 2417  
Db 19970 ACCTCCCCAGCTGAATC 19951

## RESULT 10

US-09-621-976-2835  
; Sequence 2835, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2835  
; LENGTH: 400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 99..308  
; NAME/KEY: misc\_feature  
; LOCATION: 21  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-2835

Query Match 0.5%; Score 19; DB 4; Length 400;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3573 AGCGGTGCTGAAGTCT 3591  
Db 333 AGCGGTGCTGAAGTCT 351

## RESULT 11

US-09-949-016-19079  
; Sequence 19079, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19079  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-19079

Query Match 0.5%; Score 19; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1468 TTCTTCACCTCACTCTCC 1486  
Db 29 TTCTTCACCTCACTCTCC 47

RESULT 12  
US-09-949-016-19080  
; Sequence 19080, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19080  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-19080

Query Match 0.5%; Score 19; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1468 TTCTTCACCTCACTCTCC 1486  
Db 41 TTCTTCACCTCACTCTCC 59

RESULT 13  
US-09-949-016-36893  
; Sequence 36893, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.



```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36893
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36893

Query Match          0.5%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1468 TTCTTCACCTCAGTCTCC 1486
Db      29  TTCTTCACCTCAGTCTCC 47

RESULT 14
US-09-949-016-36894
; Sequence 36894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36894
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36894

Query Match          0.5%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1468 TTCTTCACCTCAGTCTCC 1486
Db      41  TTCTTCACCTCAGTCTCC 59

RESULT 15
US-09-949-016-58788
; Sequence 58788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58788
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58788

Query Match          0.5%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2416 TCTGTGAGCTGGCAGGAA 2434
Db      147  TCTGTGAGCTGGCAGGAA 165
```

Search completed: July 19, 2005, 11:19:54  
Job time : 637 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2005, 02:27:16 ; Search time 2592 Seconds

(without alignments)  
9876.191 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atgagagagagaacatcgct.....tcgaggggctgtgtctttaa 4035

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 7173243 seqs, 3172129809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4035	100.0	4035	US-10-124-498-5	Sequence 5, Appli
2	4035	100.0	4035	US-10-066-521-5	Sequence 5, Appli
3	2917	72.3	3489	US-10-416-642-3	Sequence 3, Appli
4	2917	72.3	3926	US-10-216-645-1	Sequence 1, Appli
5	2862	70.9	3830	US-10-216-645-3	Sequence 3, Appli
6	2801	69.4	3226	US-10-092-900A-347	Sequence 347, App
7	2713	67.2	3885	US-10-860-761-3	Sequence 3, Appli

8	2713	67.2	3900	US-10-399-443-23	Sequence 23, Appli
9	2713	67.2	3900	US-10-677-943-23	Sequence 23, Appli
10	1102	27.3	1157	US-10-399-443-1	Sequence 1, Appli
11	1102	27.3	1157	US-10-677-943-1	Sequence 1, Appli
12	657	16.3	1075	US-10-399-443-3	Sequence 3, Appli
13	657	16.3	1075	US-10-677-943-3	Sequence 3, Appli
14	174	4.3	2099	US-10-027-632-258159	Sequence 258159,
15	174	4.3	2099	US-10-027-632-258159	Sequence 258159,
16	24	0.6	24	US-10-399-443-13	Sequence 13, Appli
17	24	0.6	24	US-10-399-443-14	Sequence 14, Appli
18	24	0.6	24	US-10-092-900A-750	Sequence 750, App
19	24	0.6	24	US-10-677-943-13	Sequence 13, Appli
20	24	0.6	24	US-10-677-943-14	Sequence 14, Appli
21	22	0.5	22	US-10-399-443-7	Sequence 7, Appli
22	22	0.5	22	US-10-677-943-7	Sequence 7, Appli
23	22	0.5	22	US-10-399-443-9	Sequence 9, Appli
24	22	0.5	26	US-10-677-943-9	Sequence 9, Appli
25	22	0.5	27	US-10-399-443-17	Sequence 17, Appli
26	22	0.5	27	US-10-677-943-17	Sequence 17, Appli
27	22	0.5	1036	US-10-739-930-4220	Sequence 4220, Ap
28	22	0.5	3447	US-10-399-443-5	Sequence 5, Appli
29	22	0.5	3447	US-10-677-943-5	Sequence 5, Appli
30	21	0.5	21	US-10-216-645-6	Sequence 6, Appli
31	21	0.5	21	US-10-216-645-10	Sequence 10, Appli
32	21	0.5	21	US-10-399-443-8	Sequence 8, Appli
33	21	0.5	21	US-10-399-443-19	Sequence 19, Appli
34	21	0.5	21	US-10-399-443-20	Sequence 20, Appli
35	21	0.5	21	US-10-092-900A-749	Sequence 749, App
36	21	0.5	21	US-10-092-900A-751	Sequence 751, App
37	21	0.5	21	US-10-677-943-8	Sequence 8, Appli
38	21	0.5	21	US-10-677-943-19	Sequence 19, Appli
39	21	0.5	21	US-10-677-943-20	Sequence 20, Appli
40	21	0.5	21	US-10-677-943-11	Sequence 11, Appli
41	21	0.5	22	US-10-399-443-16	Sequence 16, Appli
42	21	0.5	22	US-10-677-943-16	Sequence 16, Appli
43	21	0.5	476	US-09-918-995-62036	Sequence 62036, A
44	21	0.5	483	US-09-918-995-20036	Sequence 20036, A
45	21	0.5	495	US-10-723-860-2423	Sequence 2423, Ap

#### ALIGNMENTS

RESULT 1  
US-10-124-498-5  
; Sequence 5, Application US/10124498  
; Publication No. US20030017983A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; APPLICANT: Wang, Weiye  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-367001  
; CURRENT APPLICATION NUMBER: US/10/124, 498  
; CURRENT FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 10/066, 521  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/318, 645  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/265, 231  
; PRIOR FILING DATE: 2001-01-31  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 4035  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(4032)  
; US-10-124-498-5

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGCAAAATGCTCACTTTTTCAGCTACGGGCTGCAGATGGTGTCTATAG 60  
DB 1 ATGGAAGAGCAAAATGCTCACTTTTTCAGCTACGGGCTGCAGATGGTGTCTATAG 60

QY 61 CTAGCAAGGAAGATTTTCAACATTCAGAAATTACTAAAGAAATCTTCAAGATC 120  
DB 61 CTAGCAAGGAAGATTTTCAACATTCAGAAATTACTAAAGAAATCTTCAAGATC 120

QY 121 ACCACATGCTCTATCCACAGTTTGAATTCAGAAATGCCAAGTGAATGTCTGGCACTC 180  
DB 121 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCCAAGTGAATGTCTGGCACTC 180

QY 181 CTCTTGCAATGATATTATAGAGCATGCTGGCTGGGCTACGTCATTAGCATCTTTGAA 240  
DB 181 CTCTTGCAATGATATTATAGAGCATGCTGGCTGGGCTACGTCATTAGCATCTTTGAA 240

QY 241 AACATGAACCTGCCAACCTCTCCGAGAAAGCAAGGATGACATGAAAAAATTTTCAAA 300  
DB 241 AACATGAACCTGCCAACCTCTCCGAGAAAGCAAGGATGACATGAAAAAATTTTCAAA 300

QY 301 GCTATGAAACAAGAGTGCACAGCAGACAGACAGAAATTTTCAAGACT 360  
DB 301 GCTATGAAACAAGAGTGCACAGCAGACAGACAGAAATTTTCAAGACT 360

QY 361 ATGGAACAAGAGTGCACAGCAGACAGACAGAAACAAGACATGAGAGTGCACA 420  
DB 361 ATGGAACAAGAGTGCACAGCAGACAGACAGAAACAAGACATGAGAGTGCACA 420

QY 421 TGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGTACGTCATAGT 480  
DB 421 TGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGTACGTCATAGT 480

QY 481 TTTGAAAAACATGCTGCTGACCTGGACCGAAATGCAAAAGTGGCTGTGATTTCA 540  
DB 481 TTTGAAAAACATGCTGCTGACCTGGACCGAAATGCAAAAGTGGCTGTGATTTCA 540

QY 541 GACCGGTGGGCTTCCGGCTTCGACAGGTGTTTCGACAGAAAGTCAAGAAATTTGGAAA 600  
DB 541 GACCGGTGGGCTTCCGGCTTCGACAGGTGTTTCGACAGAAAGTCAAGAAATTTGGAAA 600

QY 601 TCGGCTACAGCAAGAGATGCTGCTGGGCTGGGCTGCAAGGTGACCTTACCAAGGAATG 660  
DB 601 TCGGCTACAGCAAGAGATGCTGCTGGGCTGGGCTGCAAGGTGACCTTACCAAGGAATG 660

QY 661 TTTCTCTACGTCTTCTCTCCCGTTAGAGAGATGACCGGAAGAGACAGAGTGC 720  
DB 661 TTTCTCTACGTCTTCTCTCCCGTTAGAGAGATGACCGGAAGAGACAGAGTGC 720

QY 721 ACAAGATTCATCTCAGAGAGTGCACAGCTTCCAGGCTCCGATGACGAATCATATGCC 780  
DB 721 ACAAGATTCATCTCAGAGAGTGCACAGCTTCCAGGCTCCGATGACGAATCATATGCC 780

QY 781 CGAACCAAGAGGCTGTTGTTATCACTTGAAGGTTTCAGATGACCTGGGCTGTCTCTCAAC 840  
DB 781 CGAACCAAGAGGCTGTTGTTATCACTTGAAGGTTTCAGATGACCTGGGCTGTCTCTCAAC 840

QY 841 AATGACACAAGAGCTTCGAAAGACTGGCTGAGAGAGCTCCCTTCAACCTCAAGC 900  
DB 841 AATGACACAAGAGCTTCGAAAGACTGGCTGAGAGAGCTCCCTTCAACCTCAAGC 900

QY 901 AGTCTGAGAGAGTCTGCTGCTGAGATCTTCTGATGCTCAAGCTGAGAGAGCTG 960  
DB 901 AGTCTGAGAGAGTCTGCTGCTGAGATCTTCTGATGCTCAAGCTGAGAGAGCTG 960

QY 961 GGCACAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTATGAGAGATC 1020  
DB 961 GGCACAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTATGAGAGATC 1020

QY 1021 TCCGGGAAACAAGAAATCCACTGTGCTCTTGAAGCGGAGATTGGTGAACATCAAGACA 1080

DB 1021 TCCGGGAAACAAGAAATCCACTGTGCTCTTGAAGCGGAGATTGGTGAACATCAAGACA 1080

QY 1081 CAAGGTTGCTGCTGATCAATGAAACACCGTGAAGTCTCCAGCAGTGCAGAGTCCGCC 1140  
DB 1081 CAAGGTTGCTGCTGATCAATGAAACACCGTGAAGTCTCCAGCAGTGCAGAGTCCGCC 1140

QY 1141 GTGGGCTCTCATCTGAGTGGCCCTGACAGTGCAGAGACGATGGGAGAGAGTCCGCC 1200  
DB 1141 GTGGGCTCTCATCTGAGTGGCCCTGACAGTGCAGAGACGATGGGAGAGAGTCCGCC 1200

QY 1201 CCTTCAACAACAAGCTCAAGGCTGACGCGCTTTGTGTTTCAATCAAGTCAACCT 1260  
DB 1201 CCTTCAACAACAAGCTCAAGGCTGACGCGCTTTGTGTTTCAATCAAGTCAACCT 1260

QY 1261 CGAGGCTGCTGCTGAGTGGCTGCTCAATCTGAGAGAAAGATTTGCTGAAGCTTCTGC 1320  
DB 1261 CGAGGCTGCTGCTGAGTGGCTGCTCAATCTGAGAGAAAGATTTGCTGAAGCTTCTGC 1320

QY 1321 CGTATGGCTGAGAGAGTGGGATAGGAAGTCAAGTGTGATGAGTGAAGACCTCATG 1380  
DB 1321 CGTATGGCTGAGAGAGTGGGATAGGAAGTCAAGTGTGATGAGTGAAGACCTCATG 1380

QY 1381 GTTCAAGAGCTCGGAGAGTCTGAGCTCGTGTCTGTTTCAATGAACATCTTTCTCCA 1440  
DB 1381 GTTCAAGAGCTCGGAGAGTCTGAGCTCGTGTCTGTTTCAATGAACATCTTTCTCCA 1440

QY 1441 GACAGCCACTGTGAGAGTACTACACTTTTCCACTCAAGTCCAGAGATTTGTGCC 1500  
DB 1441 GACAGCCACTGTGAGAGTACTACACTTTTCCACTCAAGTCCAGAGATTTGTGCC 1500

QY 1501 GCTTTGACTACAGTATTAGAGAGGCTGGAATTCAGAGCAGCTCTGCTCTGTAGCTT 1560  
DB 1501 GCTTTGACTACAGTATTAGAGAGGCTGGAATTCAGAGCAGCTCTGCTCTGTAGCTT 1560

QY 1561 GAGAAACAAGAGTCCATGAGACTTAAACAGCAGGCTTCCATATCCACTGCTTTGG 1620  
DB 1561 GAGAAACAAGAGTCCATGAGACTTAAACAGCAGGCTTCCATATCCACTGCTTTGG 1620

QY 1621 ATGAAGGCTTTCTTTTGGCTCTGCTGAGAGGAGACCTTAAGAGAGCCACTGAGAGTCTG 1680  
DB 1621 ATGAAGGCTTTCTTTTGGCTCTGCTGAGAGGAGACCTTAAGAGAGCCACTGAGAGTCTG 1680

QY 1681 CTGGGCTGTCCCGTTCCTGGGGGTGAGAGAGCTTCTGCACTGGGCTCTCTGTG 1740  
DB 1681 CTGGGCTGTCCCGTTCCTGGGGGTGAGAGAGCTTCTGCACTGGGCTCTCTGTG 1740

QY 1741 GGTGAGAGCTTAATGCCACCAACCAGAGACACCTTGAAGGCTTCACTGTCTTTTC 1800  
DB 1741 GGTGAGAGCTTAATGCCACCAACCAGAGACACCTTGAAGGCTTCACTGTCTTTTC 1800

QY 1801 GAGACTCAAGACAAGAGTGTGCTTGGCTTAAACAGCTTCCAAAGAGTGTGCTT 1860  
DB 1801 GAGACTCAAGACAAGAGTGTGCTTGGCTTAAACAGCTTCCAAAGAGTGTGCTT 1860

QY 1861 CCGATTAACAAGAGCTGAGCTTGAATGACATCTTCTTCTGCTCCAGACATGCTGAT 1920  
DB 1861 CCGATTAACAAGAGCTGAGCTTGAATGACATCTTCTTCTGCTCCAGACATGCTGAT 1920

QY 1921 TTGGGAAAAATTCGGGTGATGTTCAAGGATCTTCCCAAGATGAGTCCGCTGAGGCA 1980  
DB 1921 TTGGGAAAAATTCGGGTGATGTTCAAGGATCTTCCCAAGATGAGTCCGCTGAGGCA 1980

QY 1981 TGTCTGTGCTCTCTATGATGTCGGGATTAAGACCTCATTTGAGAGCAGTGGGAAGAT 2040  
DB 1981 TGTCTGTGCTCTCTATGATGTCGGGATTAAGACCTCATTTGAGAGCAGTGGGAAGAT 2040

QY 2041 TTTGCTTCATGCTTTGGACCAACCAACCTGAGAGCTGAGCTGAGAGAGATC 2100  
DB 2041 TTTGCTTCATGCTTTGGACCAACCAACCTGAGAGCTGAGCTGAGAGAGATC 2100

QY 2101 CTGACAGAGGAGCATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160

Db 2101 CTGACAGACGGGCGCATGAAGACCTGTGTGCCAAGCTGAGGACCTCCACCTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTTGAAGATGACAGATTATCCCTGTGTGAGAGACCTCTGGAATAC 2220  
Db 2161 CAGACCTGATGTTTGAAGATGACAGATTATCCCTGTGTGAGAGACCTCTGGAATAC 2220  
Qy 2221 GTCATGGCCAAACCGTAACTTAAGATCCCTCAACTTGGAGGACCCACCTGAAGAAAG 2280  
Db 2221 GTCATGGCCAAACCGTAACTTAAGATCCCTCAACTTGGAGGACCCACCTGAAGAAAG 2280  
Qy 2281 GATGTAAAGATGGCGTGTGAAGCTTAAACACCCAAATGTTTGTGTGAAGTCTTTGAG 2340  
Db 2281 GATGTAAAGATGGCGTGTGAAGCTTAAACACCCAAATGTTTGTGTGAAGTCTTTGAG 2340  
Qy 2341 CTGGAATGCTGTGGAATTTGACCCATGCTGTATCTTGAAGATCTCCCAATCTCTTACGAC 2400  
Db 2341 CTGGAATGCTGTGGAATTTGACCCATGCTGTATCTTGAAGATCTCCCAATCTCTTACGAC 2400  
Qy 2401 TCCCCAGGCTGAATCTCTGAGCCTGGAGGAAACAAAGTGAAGACAGACGAGAGTAAATG 2460  
Db 2401 TCCCCAGGCTGAATCTCTGAGCCTGGAGGAAACAAAGTGAAGACAGACGAGAGTAAATG 2460  
Qy 2461 CCTCTAGTATGATGCTTGAAGATCTCCAGTGGCCCTGCAAGAGCTGATCTGAAGAGAC 2520  
Db 2461 CCTCTAGTATGATGCTTGAAGATCTCCAGTGGCCCTGCAAGAGCTGATCTGAAGAGAC 2520  
Qy 2521 TGTGGCATTAAGACCCAGGCTTGGAGAGTCTGGCTCAAGCCCTGCTCAAGACCGGAGC 2580  
Db 2521 TGTGGCATTAAGACCCAGGCTTGGAGAGTCTGGCTCAAGCCCTGCTCAAGACCGGAGC 2580  
Qy 2581 TTGACACACCTGTGCTATCAACAAACAGCTGGGGAGGAAAGGTAAATCTACGTGT 2640  
Db 2581 TTGACACACCTGTGCTATCAACAAACAGCTGGGGAGGAAAGGTAAATCTACGTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATCAGTGGCCACTG 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATCAGTGGCCACTG 2700  
Qy 2701 GACAGGCTGGCTGTGTTTCTTGAATCTTGGCTTAATGAGGTAATCTCATGCTGACGAC 2760  
Db 2701 GACAGGCTGGCTGTGTTTCTTGAATCTTGGCTTAATGAGGTAATCTCATGCTGACGAC 2760  
Qy 2761 CTGAGCCTTGAAGTAAACCTGTGGAAGCAATGAGGTGAAGCTTGTGCGAGAGCTGAG 2820  
Db 2761 CTGAGCCTTGAAGTAAACCTGTGGAAGCAATGAGGTGAAGCTTGTGCGAGAGCTGAG 2820  
Qy 2821 AGAGAACCATCTTGTATCTCAAGACCTGAGATGTTGTAAGTGTATCTCAACCGCGG 2880  
Db 2821 AGAGAACCATCTTGTATCTCAAGACCTGAGATGTTGTAAGTGTATCTCAACCGCGG 2880  
Qy 2881 TGTGTGAGATCTGTCTGTGTGATCTGAGAGACGACCTGAAGAGCTGTGATCTC 2940  
Db 2881 TGTGTGAGATCTGTCTGTGTGATCTGAGAGACGACCTGAAGAGCTGTGATCTC 2940  
Qy 2941 AGGACAAATGCTGGGTGAAGCTGTGTGCTGGCTGTGAGAGGAACTGAAGCAAAAG 3000  
Db 2941 AGGACAAATGCTGGGTGAAGCTGTGTGCTGGCTGTGAGAGGAACTGAAGCAAAAG 3000  
Qy 3001 AACAGTGTCTGACGAGCTCGGCTTGAAGCATGAGCTGACTTCTGATGCTGTGAG 3060  
Db 3001 AACAGTGTCTGACGAGCTCGGCTTGAAGCATGAGCTGACTTCTGATGCTGTGAG 3060  
Qy 3061 GCACTCTCTTGGCCCTTCTGCAACCGGCACTGACAGCTTAAACCTGTGAGAAAT 3120  
Db 3061 GCACTCTCTTGGCCCTTCTGCAACCGGCACTGACAGCTTAAACCTGTGAGAAAT 3120  
Qy 3121 AACCTCAGTCCCAAGGAATGAAGCTGTGTGCTGGCTTGGCTTCCCAAGCTTAAAC 3180  
Db 3121 AACCTCAGTCCCAAGGAATGAAGCTGTGTGCTGGCTTGGCTTCCCAAGCTTAAAC 3180  
Qy 3181 TTACAGATTAATGGGCTGTGAAATGAGCATGCTGCTGCAATAAGAAAGCTGTGAG 3240  
Db 3181 TTACAGATTAATGGGCTGTGAAATGAGCATGCTGCTGCAATAAGAAAGCTGTGAG 3240

Qy 3241 GAAGTGACGCTTCTCAAGCCCGAGTGGTAATTTGACGATGTTGCAATTTTATGTA 3300  
Db 3241 GAAGTGACGCTTCTCAAGCCCGAGTGGTAATTTGACGATGTTGCAATTTTATGTA 3300  
Qy 3301 GATGACCGACCAAAATAGACTTACTTCCGGCTCCCTGAAGCGGCAATGACATG 3360  
Db 3301 GATGACCGACCAAAATAGACTTACTTCCGGCTCCCTGAAGCGGCAATGACATG 3360  
Qy 3361 GCTTGTGTGGGAGTGAACCCAGAGCAAGAAAGCTGTGTGCTTGTGCTGAGAC 3420  
Db 3361 GCTTGTGTGGGAGTGAACCCAGAGCAAGAAAGCTGTGTGCTTGTGCTGAGAC 3420  
Qy 3421 TTCAAGACAGTACACGATTTTCCCAAGTCTCTGCTGCTGCGCAACGCAATGATGCC 3480  
Db 3421 TTCAAGACAGTACACGATTTTCCCAAGTCTCTGCTGCTGCGCAACGCAATGATGCC 3480  
Qy 3481 CAGAGATTTGACAAAGTGAAGAGCTCCCGCAACCATGAGGACGGAACAA 3540  
Db 3481 CAGAGATTTGACAAAGTGAAGAGCTCCCGCAACCATGAGGACGGAACAA 3540  
Qy 3541 CAAGATTAATGTTGATGTTGATATTTCCGAGCTGTGTGAATCTGAGCTGAA 3600  
Db 3541 CAAGATTAATGTTGATGTTGATATTTCCGAGCTGTGTGAATCTGAGCTGAA 3600  
Qy 3601 GGGCTTGAATCCAAAGTGTGATATGACCAACGAGGATATGCTGTGCTCAATGAGGA 3660  
Db 3601 GGGCTTGAATCCAAAGTGTGATATGACCAACGAGGATATGCTGTGCTCAATGAGGA 3660  
Qy 3661 GAGCTGAGCTGAGGGGCTTGTGTCACAGTGTGATGACCAAGCGGCTGTGCTG 3720  
Db 3661 GAGCTGAGCTGAGGGGCTTGTGTCACAGTGTGATGACCAAGCGGCTGTGCTG 3720  
Qy 3721 CACTGGAAGCGGCTGAGCTTGAAGGCTGTGTCTTAAACAGTGTGATGACCAAGCGG 3780  
Db 3721 CACTGGAAGCGGCTGAGCTTGAAGGCTGTGTCTTAAACAGTGTGATGACCAAGCGG 3780  
Qy 3781 GTGTCTGTGCTCACTGGAGCGGCTGGGCTCGAGGAGCTTGTGTCCAAGTGTGATGAC 3840  
Db 3781 GTGTCTGTGCTCACTGGAGCGGCTGGGCTCGAGGAGCTTGTGTCCAAGTGTGATGAC 3840  
Qy 3841 CACAGCGGTGTGAGCTGTGCTGACCTGGAGCGGCTGAGGAGCTTGTGTCCAAGT 3900  
Db 3841 CACAGCGGTGTGAGCTGTGCTGACCTGGAGCGGCTGAGGAGCTTGTGTCCAAGT 3900  
Qy 3901 GCTGATGACCAAGCGGTGTGTCTGTCACTGGAGCGGCTGGGCTGAGGAGCTTGTG 3960  
Db 3901 GCTGATGACCAAGCGGTGTGTCTGTCACTGGAGCGGCTGGGCTGAGGAGCTTGTG 3960  
Qy 3961 TCCAACAGTGTGATGACCAAGCGGTGTGTCTGTCACTGGAGCGGCTGGGCTGAG 4020  
Db 3961 TCCAACAGTGTGATGACCAAGCGGTGTGTCTGTCACTGGAGCGGCTGGGCTGAG 4020  
Qy 4021 GGGCTGTGTCTTAA 4035  
Db 4021 GGGCTGTGTCTTAA 4035

RESULT 2  
US-10-066-521-5  
; Sequence 5, Application US/1006521  
; Publication No. US20030027757A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; APPLICANT: Wang, Weiye  
; APPLICANT: Blatcher, Maria  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
; FILE REFERENCE: 0734-334001  
; CURRENT APPLICATION NUMBER: US/10/066, 521  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 60/318, 645  
; PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/265,231  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 4035  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(4032)  
US-10-066-521-5.

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGCAAAATGCTCACTTTTCAAGTACGGGCTGCAATGCTGCTATGAG 60  
DB 1 ATGGAAGAGCAAAATGCTCACTTTTCAAGTACGGGCTGCAATGCTGCTATGAG 60  
QY 61 CTAGCAAGAGAAATTTTCAAGATTTCAAGAAATTTTCAAGAAATTTTCAAGAAATG 120  
DB 61 CTAGCAAGAGAAATTTTCAAGATTTCAAGAAATTTCAAGAAATTTTCAAGAAATG 120  
QY 121 ACCAATGCTCTATTCACAGATTGGAATCGAAGATGCAAGTGAATGCTGAGCACTC 180  
DB 121 ACCAATGCTCTATTCACAGATTGGAATCGAAGATGCAAGTGAATGCTGAGCACTC 180  
QY 181 CTCTTGATGATATTTATGAGATCGCTGCTGGGCTGACGTCATTAGCATCTTTGAA 240  
DB 181 CTCTTGATGATATTTATGAGATCGCTGCTGGGCTGACGTCATTAGCATCTTTGAA 240  
QY 241 AACATGAATCTGCGCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAAATTTCAAA 300  
DB 241 AACATGAATCTGCGCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAAATTTCAAA 300  
QY 301 GCTATGAAACAAGAGTGGCCACAGCAGACAGACAGAAACAAGAAATTTTCAAGACT 360  
DB 301 GCTATGAAACAAGAGTGGCCACAGCAGACAGACAGAAACAAGAAATTTTCAAGACT 360  
QY 361 ATGGAACAAGAGTGGCCACAGCAGACAGACAGAAACAAGAAACAAGAGTGGACCA 420  
DB 361 ATGGAACAAGAGTGGCCACAGCAGACAGACAGAAACAAGAGCAATGAGTGGACCA 420  
QY 421 TGGGACTCAAGAGTCAAGTATGACCAATTCGCTGAGAGAGAGATGTAAGTCTGTA 480  
DB 421 TGGGACTCAAGAGTCAAGTATGACCAATTCGCTGAGAGAGAGATGTAAGTCTGTA 480  
QY 481 TTTGAAAACATGCTGCTGACCTGACCGGAAATGCAAACTGCTGCTTTGATTCA 540  
DB 481 TTTGAAAACATGCTGCTGACCTGACCGGAAATGCAAACTGCTGCTTTGATTCA 540  
QY 541 GACCGGTGGGCTTCGCGCTCGCAAGTGGTTCGACGGAAGTCAAGAAATTTGGA 600  
DB 541 GACCGGTGGGCTTCGCGCTCGCAAGTGGTTCGACGGAAGTCAAGAAATTTGGA 600  
QY 601 TCGGCTCTAGCCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 601 TCGGCTCTAGCCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 661 TTTCTCTAGCTCTTCT 720  
DB 661 TTTCTCTAGCTCTTCT 720  
QY 721 ACAAGTTCATCTCAGGGAAGTGGCAGACTCCAGGCTCCGCTGACGGAATCATGTC 780  
DB 721 ACAAGTTCATCTCAGGGAAGTGGCAGACTCCAGGCTCCGCTGACGGAATCATGTC 780  
QY 781 CGACCAAGAAAGCTGTTGTTATCATTTGACGTTTCAATGACCTGAGCTCTGCTCAAC 840  
DB 781 CGACCAAGAAAGCTGTTGTTATCATTTGACGTTTCAATGACCTGAGCTCTGCTCAAC 840

QY 841 AATGACAAAGCTCTGCAAGATCTGGGCTGAGAAAGAGCTCCGTTCACTCATACGC 900  
DB 841 AATGACAAAGCTCTGCAAGATCTGGGCTGAGAAAGAGCTCCGTTCACTCATACGC 900  
QY 901 AGTCTGCTGAGAAAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
DB 901 AGTCTGCTGAGAAAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
QY 961 GGCACAGAAAGCTCAAGTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 961 GGCACAGAAAGCTCAAGTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1021 TCCGGGAAACAAGAAATCACTTGTCTTGAAGCGCGGATGCTGAGCATCAGAGACA 1080  
DB 1021 TCCGGGAAACAAGAAATCACTTGTCTTGAAGCGCGGATGCTGAGCATCAGAGACA 1080  
QY 1081 CAAGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1081 CAAGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 GTGGGCTCTCATCTGCGTGGCCCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
DB 1141 GTGGGCTCTCATCTGCGTGGCCCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
QY 1201 CCTTTCAACCAAGCTCAAGGCTGACAGCGCTTTTGTGTTCAATCAAGCTCACCCCT 1260  
DB 1201 CCTTTCAACCAAGCTCAAGGCTGACAGCGCTTTTGTGTTCAATCAAGCTCACCCCT 1260  
QY 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 CGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
DB 1321 CGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
QY 1381 GTTCAAGAGCTCGGGGAGTCTGAGCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
DB 1381 GTTCAAGAGCTCGGGGAGTCTGAGCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
QY 1441 GACAGCACTGCTGAGAGTACTACCTTTTCACTCACTCACTCACTCACTCACTCACT 1500  
DB 1441 GACAGCACTGCTGAGAGTACTACCTTTTCACTCACTCACTCACTCACTCACTCACT 1500  
QY 1501 GCTTGTACTACGCTTGAAGAGGCTGGAATGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
DB 1501 GCTTGTACTACGCTTGAAGAGGCTGGAATGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
QY 1561 GAGAAACAAGAGTCCATGAGACTTAAACAGCAGGCTTCAATCACTGCTGCTGCTG 1620  
DB 1561 GAGAAACAAGAGTCCATGAGACTTAAACAGCAGGCTTCAATCACTGCTGCTGCTG 1620  
QY 1621 ATGAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
DB 1621 ATGAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
DB 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
QY 1741 GGTGAGAGCTTAATGCAACCCAGAGAGACCTTGAAGGCTTCACTGCTGCTGCTGCT 1800  
DB 1741 GGTGAGAGCTTAATGCAACCCAGAGAGACCTTGAAGGCTTCACTGCTGCTGCTGCT 1800  
QY 1801 GAGACTCAAGCAAAAGATTGTTGCTTGGCATTTAAACAGCTTCAAGAAATGCTGCT 1860  
DB 1801 GAGACTCAAGCAAAAGATTGTTGCTTGGCATTTAAACAGCTTCAAGAAATGCTGCT 1860  
QY 1861 CCGATTAAACAAGAACTGGAATTGATGACATTTCTGCTGCTGCTGCTGCTGCTGCTG 1920  
DB 1861 CCGATTAAACAAGAACTGGAATTGATGACATTTCTGCTGCTGCTGCTGCTGCTGCTG 1920  
QY 1921 TTGCGAAAAATTCGGGTGATGTCAAAGGATCTTCCAAAGATGAGTCCGCTGAGGCA 1980



Db 1921 TTGGCGAAAAATTCGGGTGATGTCAAGGATCTTCCAGAGATGATGATCGCTGAGCA 1980  
Qy 1981 TGTCTGTGTGTCCTCTATGATGTGGGATTAAGACCTCTATGAGAGCATGGGAGAT 2040  
Db 1981 TGTCTGTGTGTCCTCTATGATGTGGGATTAAGACCTCTATGAGAGCATGGGAGAT 2040  
Qy 2041 TGTCTGTGTGTCCTCTATGATGTGGGATTAAGACCTCTATGAGAGCATGGGAGAT 2100  
Db 2041 TGTCTGTGTGTCCTCTATGATGTGGGATTAAGACCTCTATGAGAGCATGGGAGAT 2100  
Qy 2101 CTGACAGAGCGGGCATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGGCATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGTGATTTTGAATGACAGATTAACCTGTGTGAGACCTGTGAGATC 2220  
Db 2161 CAGACCTGTGATTTTGAATGACAGATTAACCTGTGTGAGACCTGTGAGATC 2220  
Qy 2221 GTCATGGGCAACCGTAACCTAGATCCCTCAACTGGGAGGACCACTGAAAGAGAG 2280  
Db 2221 GTCATGGGCAACCGTAACCTAGATCCCTCAACTGGGAGGACCACTGAAAGAGAG 2280  
Qy 2281 GATGTAAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTGTGTGAGCTTTGAGG 2340  
Db 2281 GATGTAAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTGTGTGAGCTTTGAGG 2340  
Qy 2341 CTGGAATGTGTGATGATGACCAATGCTGTTAAGATGATCCCAATCTTACGACC 2400  
Db 2341 CTGGAATGTGTGATGATGACCAATGCTGTTAAGATGATCCCAATCTTACGACC 2400  
Qy 2401 TCCCCAGGCTTAAATCTGTGAGCTGTGAGAGAAACAGATGACAGACAGAGATATG 2460  
Db 2401 TCCCCAGGCTTAAATCTGTGAGCTGTGAGAGAAACAGATGACAGACAGAGATATG 2460  
Qy 2461 CCTCTAGATGACCTGTGAGATCTCCAGTGCGCCCTGACAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTAGATGACCTGTGAGATCTCCAGTGCGCCCTGACAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGACATCAAGCCAGGGTGTGCAAGATGAGCTGACCCCTGACAGACCGGAGAC 2580  
Db 2521 TGTGACATCAAGCCAGGGTGTGCAAGATGAGCTGACCCCTGACAGACCGGAGAC 2580  
Qy 2581 TTGACACACCTGTGCTATCCAAACAAGCTGTGGGAAACAGAGGTGTAATCTACTGTGT 2640  
Db 2581 TTGACACACCTGTGCTATCCAAACAAGCTGTGGGAAACAGAGGTGTAATCTACTGTGT 2640  
Qy 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGATCACTGACCTG 2700  
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGATCACTGACCTG 2700  
Qy 2701 GACACGGCTGTGTGTGTTTCTTCACTTGGCTTAAGGATCACTATGCTGACGAC 2760  
Db 2701 GACACGGCTGTGTGTGTTTCTTCACTTGGCTTAAGGATCACTATGCTGACGAC 2760  
Qy 2761 CTGAGCTTGAATGAACCTGTGGAAGACATGCGTGAAGCTTCTGTGCGAGGT CATG 2820  
Db 2761 CTGAGCTTGAATGAACCTGTGGAAGACATGCGTGAAGCTTCTGTGCGAGGT CATG 2820  
Qy 2821 AGAGAACCATCTTGCATCTCAAGACCTGGAGTTGTAAGTGCATCTCAACCGCGCG 2880  
Db 2821 AGAGAACCATCTTGCATCTCAAGACCTGGAGTTGTAAGTGCATCTCAACCGCGCG 2880  
Qy 2881 TGTCTGTGAGATCTGTCTGTGTGATCTGAGAGACAGACCTGAGAGCTGTGATCTC 2940  
Db 2881 TGTCTGTGAGATCTGTCTGTGTGATCTGAGAGACAGACCTGAGAGCTGTGATCTC 2940  
Qy 2941 ACGGACATGCTGTGTGAGACGCTGTGGGTGTGCTGTGTGAGAGGAGCTGAAGAGAG 3000  
Db 2941 ACGGACATGCTGTGTGAGACGCTGTGGGTGTGCTGTGTGAGAGGAGCTGAAGAGAG 3000  
Qy 3001 AACAGTGTCTGACAGAGCTGGGTGTAAGGAGCATGTGACTGATCTGATGCTGTGAG 3060  
Db 3001 AACAGTGTCTGACAGAGCTGGGTGTAAGGAGCATGTGACTGATCTGATGCTGTGAG 3060

Db 3001 AACAGTGTCTGACAGAGCTGGGTGTAAGGAGCATGTGACTGATCTGATGCTGTGAG 3060  
Qy 3061 GCATCTTCTTGGGCTTCTCTGCAACCGGATCTGACCACTTAACCTGTGTGAGAT 3120  
Db 3061 GCATCTTCTTGGGCTTCTCTGCAACCGGATCTGACCACTTAACCTGTGTGAGAT 3120  
Qy 3121 AACTCAGTCCCAAGAAATGATGAAGCTGTGTGAGGCTTGTGCTGTGCTGAC 3180  
Db 3121 AACTCAGTCCCAAGAAATGATGAAGCTGTGTGAGGCTTGTGCTGTGCTGAC 3180  
Qy 3181 TTACAGATTAATGTGGCTGTGAAATGGCAGTACCTGTGCAATTAAGAGCTGTGAG 3240  
Db 3181 TTACAGATTAATGTGGCTGTGAAATGGCAGTACCTGTGCAATTAAGAGCTGTGAG 3240  
Qy 3241 GAATGACGCTACTCAACCCCGAGTGTGATTTGACGATGTGCTTTTGTGATGA 3300  
Db 3241 GAATGACGCTACTCAACCCCGAGTGTGATTTGACGATGTGCTTTTGTGATGA 3300  
Qy 3301 GATGACGACCAAAATGAGACTTAATCTTCCGGCTCCCTGAAAGCCGGGCAATGCTATG 3360  
Db 3301 GATGACGACCAAAATGAGACTTAATCTTCCGGCTCCCTGAAAGCCGGGCAATGCTATG 3360  
Qy 3361 GCTTGTGTGGGAGTGAACCAAGAGAGAAAGCCTGTGTGCTTCTGTGAGAC 3420  
Db 3361 GCTTGTGTGGGAGTGAACCAAGAGAGAAAGCCTGTGTGCTTCTGTGAGAC 3420  
Qy 3421 TTCAAGACAGTACAGATTTTCCAGATCTCTGTGCTGTGCAACGCAATGTGTGATCC 3480  
Db 3421 TTCAAGACAGTACAGATTTTCCAGATCTCTGTGCTGTGCAACGCAATGTGTGATCC 3480  
Qy 3481 CAGAGAGTTGACAACTGTGACAGAGCTCCCGCAACCAAGGAGAGAGAGAGAGAGAG 3540  
Db 3481 CAGAGAGTTGACAACTGTGACAGAGCTCCCGCAACCAAGGAGAGAGAGAGAGAGAG 3540  
Qy 3541 CAAGATTAATGTTGAGTGTGATATTCGAGAGCTGTGCTGAACTGTGAGCTGAA 3600  
Db 3541 CAAGATTAATGTTGAGTGTGATATTCGAGAGCTGTGCTGAACTGTGAGCTGAA 3600  
Qy 3601 GGGCTTGTGATCAACAGTGTGATGATCAACAGAGGATGAGCTGTGCTGATGAGAG 3660  
Db 3601 GGGCTTGTGATCAACAGTGTGATGATCAACAGAGGATGAGCTGTGCTGATGAGAG 3660  
Qy 3661 GAGCTGACCTGAGGGGCTGTGTGCAACAGTGTGATGACCAAGCGGTGTGTCTGTGT 3720  
Db 3661 GAGCTGACCTGAGGGGCTGTGTGCAACAGTGTGATGACCAAGCGGTGTGTCTGTGT 3720  
Qy 3721 CACTGAGAGCGGCTGTGAGGAGCTGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780  
Db 3721 CACTGAGAGCGGCTGTGAGGAGCTGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780  
Qy 3781 GTGTCTGTGTCACTGAGAGCGGCTGTGAGGAGCTGTGTCTTAAACAGTGTGATGAC 3840  
Db 3781 GTGTCTGTGTCACTGAGAGCGGCTGTGAGGAGCTGTGTCTTAAACAGTGTGATGAC 3840  
Qy 3841 CACAGCGGT 3900  
Db 3841 CACAGCGGT 3900  
Qy 3901 GGTGATGACCAAGCGGT 3960  
Db 3901 GGTGATGACCAAGCGGT 3960  
Qy 3961 TCCAAACAGT 4020  
Db 3961 TCCAAACAGT 4020  
Qy 4021 GGGCTGTGTGTCTTAA 4035  
Db 4021 GGGCTGTGTGTCTTAA 4035

RESULT 3  
US-10-416-642-3

Sequence 3, Application US/10416642  
Publication No. US20040043452A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: RANIKUMAR, Jayalaxmi  
APPLICANT: ARVIZU, Chandra  
TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0842 PCT  
CURRENT APPLICATION NUMBER: US/10/416,642  
PRIOR FILING DATE: 2003-05-13  
PRIOR APPLICATION NUMBER: 60/249,407  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PERL Program  
SEQ ID NO 3  
LENGTH: 3489  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CB1  
US-10-416-642-3

Query Match 72.3%; Score 2917; DB 18; Length 3489;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 290 AAATTTCAAGCTATGGAACAGAGAGGTGCGACAGCAGCAGAGCAGAGAAACAGAA 349  
DB 452 AAATTTCAAGCTATGGAACAGAGAGGTGCGACAGCAGCAGAGCAGAGAAACAGAA 511  
QY 350 TTTTCAAGCTATGGAACAGAGAGGTGCGACAGCAGCAGAGCAGAGAAACAGAGACATG 409  
DB 512 TTTTCAAGCTATGGAACAGAGAGGTGCGACAGCAGCAGAGCAGAGAAACAGAGACATG 571  
QY 410 GAGGTGACATGAGGAGCTACAGAGTCACTGATGACCAAAATTCGCTGAGAGAGAGATG 469  
DB 572 GAGGTGACATGAGGAGCTACAGAGTCACTGATGACCAAAATTCGCTGAGAGAGAGATG 631  
QY 470 TACGTCGATGTTTGAACACCTGCTGCTGACCTGCGCGGAAATGAAACGTTGCTGCTG 529  
DB 632 TACGTCGATGTTTGAACACCTGCTGCTGACCTGCGCGGAAATGAAACGTTGCTGCTG 691  
QY 530 CTTTGAATTCAGACCGGTGAGGCTTCGCGCTCGACAGGTGATCTGACAGGAAAGTCAAG 589  
DB 692 CTTTGAATTCAGACCGGTGAGGCTTCGCGCTCGACAGGTGATCTGACAGGAAAGTCAAG 751  
QY 590 GAATTTGGAATTCGCTCTGAGCAGAGAGATGCTGCTGCTGCTGCGCGCAAGGTGATCT 649  
DB 752 GAATTTGGAATTCGCTCTGAGCAGAGAGATGCTGCTGCTGCTGCGCGCAAGGTGATCT 811  
QY 650 ACCGCGGATGTTTCCCTACGCTCTTCCCTCCCGTTGAGAGATGAGAGGAAAGAG 709  
DB 812 ACCGCGGATGTTTCCCTACGCTCTTCCCTCCCGTTGAGAGATGAGAGGAAAGAG 871  
QY 710 AGAGCAGGTGACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTGACG 769  
DB 872 AGAGCAGGTGACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTGACG 931  
QY 770 AGATCATGTCCCGACCAAGAGGCTGTTGTTTCATTTGACGCTTTCATGACCTGAGCT 829  
DB 932 AGATCATGTCCCGACCAAGAGGCTGTTGTTTCATTTGACGCTTTCATGACCTGAGCT 991  
QY 830 CTGTCCTCAACAGATGACCAAGGCTCTGCAAGAGATGAGAGAGAGAGCTCCGTTCA 889  
DB 992 CTGTCCTCAACAGATGACCAAGGCTCTGCAAGAGATGAGAGAGAGAGCTCCGTTCA 1051  
QY 890 CCTCTATAGCAGAGTGTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATCTGACCG 949  
DB 1052 CCTCTATAGCAGAGTGTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATCTGACCG 1111  
QY 950 TCAGAGCAGTGGGACAGAGAGGCTCAAGTCAAGAGTCTGCTCCCGTTACCTGTTAG 1009

DB 1112 TCAGAGCAGTGGGACAGAGAGGCTCAAGTCAAGAGTCTGCTCCCGTTACCTGTTAG 1171  
QY 1010 TTAGAGGAATTCCTCGGGGAAACAAGAAATCACTTGTCTTTGAGCGGGATTGGTAGC 1069  
DB 1172 TTAGAGGAATTCCTCGGGGAAACAAGAAATCACTTGTCTTTGAGCGGGATTGGTAGC 1231  
QY 1070 ATCAGAGACACAGAGGTGGTGGTCCATGATGAAACAAGTGAAGTCTGACCAAGTGC 1129  
DB 1232 ATCAGAGACACAGAGGTGGTGGTCCATGATGAAACAAGTGAAGTCTGACCAAGTGC 1291  
QY 1130 AGGTGCCCCCGGTGGCTCTCTCATCTGCTGGGCCCTGACAGCTGACAGAGTGGGGG 1189  
DB 1292 AGGTGCCCCCGGTGGCTCTCTCATCTGCTGGGCCCTGACAGCTGACAGAGTGGGGG 1351  
QY 1190 AGAGCGTGGCCCCCTTCAACCAAGGCTCACAGGCTGACGCGGCTTTGTGTTTATC 1249  
DB 1352 AGAGCGTGGCCCCCTTCAACCAAGGCTCACAGGCTGACGCGGCTTTGTGTTTATC 1411  
QY 1250 AGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTGCTTGA 1309  
DB 1412 AGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTGCTTGA 1471  
QY 1310 AGCGCTTCTGCGCTGATGAGCTGAGAGAGTGGAGATAGGAATCACTGTTGATGAGTG 1369  
DB 1472 AGCGCTTCTGCGCTGATGAGCTGAGAGAGTGGAGATAGGAATCACTGTTGATGAGTG 1531  
QY 1370 ACAGCTCATAGTTCAGAGACTCGGGGAGTCTGAGCTCCGCTGCTGTTTCAACATGAA 1429  
DB 1532 ACAGCTCATAGTTCAGAGACTCGGGGAGTCTGAGCTCCGCTGCTGTTTCAACATGAA 1591  
QY 1430 TCCTTCTCCAGACAGCCACTGTGAGAGATCTACACTTCTTTCACCTCACTGCTCAAG 1489  
DB 1592 TCCTTCTCCAGACAGCCACTGTGAGAGATCTACACTTCTTTCACCTCACTGCTCAAG 1651  
QY 1490 ACTTCTGTGCGCTTGTGATCACTGTTTGAAGGCTCTGGAATGAGACCAAGCTCTGCGC 1549  
DB 1652 ACTTCTGTGCGCTTGTGATCACTGTTTGAAGGCTCTGGAATGAGACCAAGCTCTGCGC 1711  
QY 1550 CTCTGTAGTGAAGAACAAAGAGTCCATGAGCTTTAAACAGGCGGCTTCCATATCC 1609  
DB 1712 CTCTGTAGTGAAGAACAAAGAGTCCATGAGCTTTAAACAGGCGGCTTCCATATCC 1771  
QY 1610 ACTGCTTTGATGATGAGCTTCTTGTGAGCTCTGAGAGAGAGAGAGAGGACAC 1669  
DB 1772 ACTGCTTTGATGATGAGCTTCTTGTGAGCTCTGAGAGAGAGAGAGAGGAGGACAC 1831  
QY 1670 TGAAGTCTGCTGCGCTGTGCTCCCTGCGGGGAGAGAGAGAGTCTGACAGTGG 1729  
DB 1832 TGAAGTCTGCTGCGCTGTGCTCCCTGCGGGGAGAGAGAGAGTCTGACAGTGG 1891  
QY 1730 TCTCTGTGAGGCTCAGACAGCTAATGACACACCCAGAGAGACACCTGGAAGCTTCC 1789  
DB 1892 TCTCTGTGAGGCTCAGACAGCTAATGACACACCCAGAGAGACACCTGGAAGCTTCC 1951  
QY 1790 ACTGCTTTTGAAGCTCAAGACAAAGAGTGTGCTTGGCACTTAAACAGCTTCCAG 1849  
DB 1952 ACTGCTTTTGAAGCTCAAGACAAAGAGTGTGCTTGGCACTTAAACAGCTTCCAG 2011  
QY 1850 AAGTGTGGCTTCGATTAACCAAGACCTGAGACTTGAAGATCTTCTTGTGCTCCAGC 1909  
DB 2012 AAGTGTGGCTTCGATTAACCAAGACCTGAGACTTGAAGATCTTCTTGTGCTCCAGC 2071  
QY 1910 ACTGTCGATATTGCGGAAATTCGAGGTGATGTCAAAGGAGATCTTCCCAAGAGATGAGT 1969  
DB 2072 ACTGTCGATATTGCGGAAATTCGAGGTGATGTCAAAGGAGATCTTCCCAAGAGATGAGT 2131  
QY 1970 CCGCTGAGGACATGTCCTGAGTCCCTCATGATGCGGATTAAGACCTCTCATTTGAGAGC 2029  
DB 2132 CCGCTGAGGACATGTCCTGAGTCCCTCATGATGCGGATTAAGACCTCTCATTTGAGAGC 2191  
QY 2030 AGTGGAAAGATTTTGTCTCCATGCTTGGCACCCACACCTGCGGAGCTGAGCTGAG 2089  
DB 2192 AGTGGAAAGATTTTGTCTCCATGCTTGGCACCCACACCTGCGGAGCTGAGCTGAG 2251

2090 GCAGGAGCATCTCTGACAGAGCGGGCCATGAAGACCTGTGTGCGAAGCTGAGGATCCCA 2149  
2252 GCAGGAGCATCTCTGACAGAGCGGGCCATGAAGACCTGTGTGCGAAGCTGAGGATCCCA 2311  
2150 CCTGCAAGATACAGACCTGTGTGAGAAATGACAGATTAACCTGTGTGAGAGAC 2209  
2312 CCTGCAAGATACAGACCTGTGTGAGAAATGACAGATTAACCTGTGTGAGAGAC 2371  
2210 TCTGAGAAATGCTCATGAGCCAACTGAATTCCTCACTTGGAGGACACCCAC 2269  
2372 TCTGAGAAATGCTCATGAGCCAACTGAATTCCTCACTTGGAGGACACCCAC 2431  
2270 TGAAGAAAGAGATGTAAAGATGGGCTGAGACCTTAAACACCCAAATGTTGTGG 2329  
2432 TGAAGAAAGAGATGTAAAGATGGGCTGAGACCTTAAACACCCAAATGTTGTGG 2491  
2330 AGTCTTTGAGGCTGAGATGCTGTGATTAACCAATGCTGTAACCTGAAGATCTCCAA 2389  
2492 AGTCTTTGAGGCTGAGATGCTGTGATTAACCAATGCTGTAACCTGAAGATCTCCAA 2551  
2390 TCTTTACGACTCTCCCAAGCTGAAATCTTGAAGCTGAGAGAAACAGGTGACAGACC 2449  
2552 TCTTTACGACTCTCCCAAGCTGAAATCTTGAAGCTGAGAGAAACAGGTGACAGACC 2611  
2450 AGGAGTAATGCTCTCATGATGCTGAGAGCTTCCAGAGGCGCCCTGAGAGAGCTGA 2509  
2612 AGGAGTAATGCTCTCATGATGCTGAGAGCTTCCAGAGGCGCCCTGAGAGAGCTGA 2671  
2510 TACTGAGAGACTGTGGCATCAAGCCAGGAGTGGCAGAGTGGCTCAGCCCTGCTCA 2569  
2672 TACTGAGAGACTGTGGCATCAAGCCAGGAGTGGCAGAGTGGCTCAGCCCTGCTCA 2731  
2570 GCAACCGAGGCTTGAACAACCTGTGCTATTCACAACAAGCTGGGAAACGAAAGTGTAA 2629  
2732 GCAACCGAGGCTTGAACAACCTGTGCTATTCACAACAAGCTGGGAAACGAAAGTGTAA 2791  
2650 ATCTACTGTGTGATTCATCAAGAGGCTTCCCACTGATAGCTGAGAGGCTGAGAGCTGAATC 2689  
2792 ATCTACTGTGTGATTCATCAAGAGGCTTCCCACTGATAGCTGAGAGGCTGAGAGCTGAATC 2851  
2690 AGTGCACCTGGAACAAGGCTGTGTGTTTCTTGACCTTGAAGGCTGAATCTCAT 2749  
2852 AGTGCACCTGGAACAAGGCTGTGTGTTTCTTGACCTTGAAGGCTGAATCTCAT 2911  
2750 GAGTGAAGCACTGAGCTTGAAGCAATCCCTGTGAAAGCAATGAGCTGTCTGT 2809  
2912 GAGTGAAGCACTGAGCTTGAAGCAATCCCTGTGAAAGCAATGAGCTGTCTGT 2971  
2810 GCGAGGTGATGAGAGAAACCATTTGTCTATCTCAAGAGCTGAGAGTGTAAAGTGTATC 2869  
2972 GCGAGGTGATGAGAGAAACCATTTGTCTATCTCAAGAGCTGAGAGTGTAAAGTGTATC 3031  
2870 TCACCGCGCGCTGTGTGAGAGTGTGTCTGTGATCTGAGAGAGAGACCACTGAAGA 2929  
3032 TCACCGCGCGCTGTGTGAGAGTGTGTCTGTGATCTGAGAGAGAGACCACTGAAGA 3091  
2930 GCTGTGATCTCAAGCAATGCTGTGAGTGAAGTGGAGTGTGCTGCTGTGCGAGGAGC 2989  
3092 GCTGTGATCTCAAGCAATGCTGTGAGTGAAGTGGAGTGTGCTGCTGCGAGGAGC 3151  
2990 TGAAGCAAAAGAAAGAGTGTGTCTGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAG 3049  
3152 TGAAGCAAAAGAAAGAGTGTGTCTGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAG 3211  
3050 ATTGTGTGAGAGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3109  
3212 ATTGTGTGAGAGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3271  
3110 TGTGTGAGAGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3169  
3272 TGTGTGAGAGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3331

3170 CCAGCTCACTTACAGATTAATGGGCTGTGAAATGAGCAGTACCTGTGCAAAATAGGA 3229  
3332 CCAGCTCACTTACAGATTAATGGGCTGTGAAATGAGCAGTACCTGTGCAAAATAGGA 3391  
3230 AGCTGTGAGAGAGAGTCACTCAAGCCCGAGTGTAAATGAGAGAGTGTGCAAT 3289  
3392 AGCTGTGAGAGAGAGTCACTCAAGCCCGAGTGTAAATGAGAGAGTGTGCAAT 3451  
3290 CTTTGAATGAATGACCG 3308  
3452 CTTTGAATGAATGACCG 3470

RESULT 4  
US-10-216-645-1  
; Sequence 1, Application US/10216645  
; Publication No. US20030125282A1  
; GENERAL INFORMATION:  
; APPLICANT: WEISS, BERTRAM  
; APPLICANT: LESTL, MONIKA  
; APPLICANT: PETERS-KOTTIG, MICHAEL  
; APPLICANT: BECKMANN, GEORG  
; TITLE OF INVENTION: HUMAN MATER PROTEINS  
; FILE REFERENCE: SCH-1910  
; CURRENT APPLICATION NUMBER: US/10/216,645  
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3926  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-216-645-1

Query Match 72.3%; Score 2917; DB 15; Length 3926;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

290 AAATTTCAAGCTATGAGAAACAAAGAGTGCACAGCAGCAGAGACAGAAACAAAGAA 349  
452 AAATTTCAAGCTATGAGAAACAAAGAGTGCACAGCAGCAGAGACAGAAACAAAGAA 511  
350 TTTCAAGCTATGAGAAACAAAGAGTGCACAGCAGCAGAGACAGAAACAAAGACATG 409  
512 TTTCAAGCTATGAGAAACAAAGAGTGCACAGCAGCAGAGACAGAAACAAAGACATG 571  
410 GAGGTGACATGAGACATGAAGAGTCACTGATGACCAATTTGCTGAGAGAGAGATG 469  
572 GAGGTGACATGAGACATGAAGAGTCACTGATGACCAATTTGCTGAGAGAGAGATG 631  
470 TACGTGATGTTTGAAGAACTGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
632 TACGTGATGTTTGAAGAACTGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 691  
530 CTTTGAATGAGACAGTGTGGGCTTCCGCTGTGACAGAGTGTCTGACGAAAGTCAAG 589  
692 CTTTGAATGAGACAGTGTGGGCTTCCGCTGTGACAGAGTGTCTGACGAAAGTCAAG 751  
590 GAATTTGGAATTCGCTCTTGAAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 649  
752 GAATTTGGAATTCGCTCTTGAAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 811  
650 ACCAGGAAATGTTCTCTCACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 709  
812 ACCAGGAAATGTTCTCTCACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871  
710 AGAGCAAGTGCAGAGAGTCTTCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 769  
872 AGAGCAAGTGCAGAGAGTCTTCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 931  
770 AGATCATGTCCCGACAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 829

Db 932 AGATCATGTCGCCGACCAAGAAAGGCTGTTGTCATCATTTGACGGTTTCGATGACCTCGGCT 991  
Qy 830 CTGTCTCAACAATATACAAAGCTCTGCAAAAGATGAGGCTGAGAAAGAGCTCCGTTCA 889  
Db 992 CTGTCTCAACAATATACAAAGCTCTGCAAAAGATGAGGCTGAGAAAGAGCTCCGTTCA 1051  
Qy 890 CCTCTCATAGCAGATGTCTGAGAGAGATCTGCTCCCTGAGATCTTCTCGATCTGCAACG 949  
Db 1052 CCTCTCATAGCAGATGTCTGAGAGAGATCTGCTCCCTGAGATCTTCTCGATCTGCAACG 1111  
Qy 950 TCAGAGAGTGTGGGCA CAGAGAACTCAAGTCAAGAGTGTGTCTTCCGTTACCTGTTAG 1009  
Db 1112 TCAGAGAGTGTGGGCA CAGAGAACTCAAGTCAAGAGTGTGTCTTCCGTTACCTGTTAG 1171  
Qy 1010 TTAGAGAAATCTCCGGGGAA CAAAGATCCATCTTCTCTGAGGCGGGATTGTTAGC 1069  
Db 1172 TTAGAGAAATCTCCGGGGAA CAAAGATCCATCTTCTCTGAGGCGGGATTGTTAGC 1231  
Qy 1070 ATCAGAAACA CAAAGGTTGCGTGCATCATGAACA CCGTAGAGCTGCTCGACAGTGC 1129  
Db 1232 ATCAGAAACA CAAAGGTTGCGTGCATCATGAACA CCGTAGAGCTGCTCGACAGTGC 1291  
Qy 1130 AGGTGCGCGGTGGGCTCTCTCATCTGCGTGCAGCTGAGAGAGCGTGTGGGG 1189  
Db 1292 AGGTGCGCGGTGGGCTCTCTCATCTGCGTGCAGCTGAGAGAGCGTGTGGGG 1351  
Qy 1190 AGAGTGTGCGCCCTTCACAA CCAAGCTCACAGGCTGCAAGCGCTTTTGTGTTATC 1249  
Db 1352 AGAGTGTGCGCCCTTCACAA CCAAGCTCACAGGCTGCAAGCGCTTTTGTGTTATC 1411  
Qy 1250 AGCTCAACCCCTGAGAGCGTGTGCGGCGCTGCTCTCAATCTGAGAGAAAGATTGCTGA 1309  
Db 1412 AGCTCAACCCCTGAGAGCGTGTGCGGCGCTGCTCTCAATCTGAGAGAAAGATTGCTGA 1471  
Qy 1310 AGCGCTTGTGCGCTGATGAGTGTGAGGAGTGTGAATAGAAAGTCAAGTGTGATGTG 1369  
Db 1472 AGCGCTTGTGCGCTGATGAGTGTGAGGAGTGTGAATAGAAAGTCAAGTGTGATGTG 1531  
Qy 1370 ACGACCTATAGTTCACAGAACTCGGGGAGCTGAGAGTCCGTCTGTTTCAANTGAACA 1429  
Db 1532 ACGACCTATAGTTCACAGAACTCGGGGAGCTGAGAGTCCGTCTGTTTCAANTGAACA 1591  
Qy 1430 TCTCTTCCAGACAGCAGCTGAGAGAGTACTACACTTCTTCCACTCAGTCTCCAG 1489  
Db 1592 TCTCTTCCAGACAGCAGCTGAGAGAGTACTACACTTCTTCCACTCAGTCTCCAG 1651  
Qy 1490 ACTTCTGTGCGCTTGTACTAGCTGTGAGAGGCTGAGAAATCGAGCAGCTCTGCGC 1549  
Db 1652 ACTTCTGTGCGCTTGTACTAGCTGTGAGAGGCTGAGAAATCGAGCAGCTCTGCGC 1711  
Qy 1550 CTCTGTAGTGAAGAA CAAAGAGTTCATGAGACTTAACAGGCAAGCTTCCATATCC 1609  
Db 1712 CTCTGTAGTGAAGAA CAAAGAGTTCATGAGACTTAACAGGCAAGCTTCCATATCC 1771  
Qy 1610 ACTGCTTTGATGAAGGCTTCTTGTGCTGTGAGCGAAGACTTAAGAGAGCCAC 1669  
Db 1772 ACTGCTTTGATGAAGGCTTCTTGTGCTGTGAGCGAAGACTTAAGAGAGCCAC 1831  
Qy 1670 TGAAGTCTGTGAGGCTGTCCCTTCCCTGAGGAGTGAACAGAAAGCTTGTGACTGG 1729  
Db 1832 TGAAGTCTGTGAGGCTGTCCCTTCCCTGAGGAGTGAACAGAAAGCTTGTGACTGG 1891  
Qy 1730 TCTCTCTGTGGGTCAAGAGCTTAATGCAACCCCGAGAGACACCTTGAAGCTTCC 1789  
Db 1892 TCTCTCTGTGGGTCAAGAGCTTAATGCAACCCCGAGAGACACCTTGAAGCTTCC 1951  
Qy 1790 ACTGCTTTTGAAGCTCAAGCAAAAGATTGTTGCTGAGCAATTAAAGCTTCCAG 1849  
Db 1952 ACTGCTTTTGAAGCTCAAGCAAAAGATTGTTGCTGAGCAATTAAAGCTTCCAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAACAGAACTGGACTTGATGACTTCTTCTGCTCCAGC 1909

Db 2012 AAGTGTGCTTCCGATTAACAGAACTGGACTTGATGACTTCTTCTGCTCCAGC 2071  
Qy 1910 ACTGTCCGTAATTTGCGGAAATTCGGGTGATGTCAAAAGGATCTTCCAAAGATAGT 1969  
Db 2072 ACTGTCCGTAATTTGCGGAAATTTGGGTGATGTCAAAAGGATCTTCCAAAGATAGT 2131  
Qy 1970 CCGCTGAGGCAATGTCTGTGTGCTCCCTATATGATGCGGATTAAGACCTCATTTAGAGAGC 2029  
Db 2132 CCGCTGAGGCAATGTCTGTGTGCTCCCTATATGATGCGGATTAAGACCTCATTTAGAGAGC 2191  
Qy 2030 AGTGGAAAGATTTCTGCTCCATGCTTGGCACACCCACACCTGCGGAGCTGGAACCTGG 2089  
Db 2192 AGTGGAAAGATTTCTGCTCCATGCTTGGCACACCCACACCTGCGGAGCTGGAACCTGG 2251  
Qy 2090 GCACAGATCTTCAGAGAGGCGGCATGAAGACCTGTGTGCGCAAGCTGAGAGCATCCA 2149  
Db 2252 GCACAGATCTTCAGAGAGGCGGCATGAAGACCTGTGTGCGCAAGCTGAGAGCATCCA 2311  
Qy 2150 CCTGCAAGATPACAGACCCCTGATGTTTGAAGATGACAGATTACCCCTGTGTGAGAGAC 2209  
Db 2312 CCTGCAAGATPACAGACCCCTGATGTTTGAAGATGACAGATTACCCCTGTGTGAGAGAC 2371  
Qy 2210 TCTGAGAAATCTCATGAGCCACCTGTAACCTTAAGATCTTCACTTGGAGAGCACCAAC 2269  
Db 2372 TCTGAGAAATCTCATGAGCCACCTGTAACCTTAAGATCTTCACTTGGAGAGCACCAAC 2431  
Qy 2270 TGAAGAAAGGATGTAAGATGCGTGAAGCTTAAACACCCAAATGTTGTGTG 2339  
Db 2432 TGAAGAAAGGATGTAAGATGCGTGAAGCTTAAACACCCAAATGTTGTGTG 2491  
Qy 2330 AGTCTTTGAGGCTGATGCTGTGTGATTTGACCCATGCTGTATCTGAAGATCTCCCAA 2389  
Db 2492 AGTCTTTGAGGCTGATGCTGTGTGATTTGACCCATGCTGTATCTGAAGATCTCCCAA 2551  
Qy 2290 TCTTTAGACCTCCCTCCAGCTGAAATCTCTGAGCTGCGAGAGAAACAAGTGCACAGC 2449  
Db 2552 TCTTTAGACCTCCCTCCAGCTGAAATCTCTGAGCTGCGAGAGAAACAAGTGCACAGC 2611  
Qy 2450 AGGAGTAATGCTCTGAGTATGCTTGAAGATCTCCAGTGGGCCCTGCAGAAAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTGAGTATGCTTGAAGATCTCCAGTGGGCCCTGCAGAAAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGAGTACACAGCCAGGATTCAGAGTCTTGAGCTCAGCCCTGCTCA 2569  
Db 2672 TACTGAGAGACTGTGAGTACACAGCCAGGATTCAGAGTCTTGAGCTCAGCCCTGCTCA 2731  
Qy 2570 GCAACCGGAGCTTGAACACCTGTGCTATTCACAA CAGCTGCGGAGAAAGAGTGTAA 2629  
Db 2732 GCAACCGGAGCTTGAACACCTGTGCTATTCACAA CAGCTGCGGAGAAAGAGTGTAA 2791  
Qy 2630 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTACTGCAAGGCTGATGTGAATC 2689  
Db 2792 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTACTGCAAGGCTGATGTGAATC 2851  
Qy 2690 AGTGCACCTGAGACACCGGCTGTGTGTTTCTTGTGACTTGTGCGCTTATGAGTAACTCAT 2749  
Db 2852 AGTGCACCTGAGACACCGGCTGTGTGTTTCTTGTGACTTGTGCGCTTATGAGTAACTCAT 2911  
Qy 2750 GGTGAGCAGACTGAGCTTATGATGAACCTGTGTGAAGACATGCGTGAAGCTTCTGT 2809  
Db 2912 GGTGAGCAGACTGAGCTTATGATGAACCTGTGTGAAGACATGCGTGAAGCTTCTGT 2971  
Qy 2810 GCGAGGTCAATGAGAGACCAATCTGTGATCTCCAGAGCCTGAGATTGATAAGTGCATC 2869  
Db 2972 GCGAGGTCAATGAGAGACCAATCTGTGATCTCCAGAGCCTGAGATTGATAAGTGCATC 3031  
Qy 2870 TCAACCGCGGCTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACCTGTAAGA 2929  
Db 3032 TCAACCGCGGCTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACCTGTAAGA 3091  
Qy 2930 GCTGTGATCTCAACGACAAATGCCCTGGGTGACAGGTGGGTTGCTGCGCTGTGCGAGGAC 2989  
Db 3092 GCTGTGATCTCAACGACAAATGCCCTGGGTGACAGGTGGGTTGCTGCGCTGTGCGAGGAC 3151

QY 2990 TGAAGCAAAAGAACAGTGTCTGACGAGACTGGGTTGAAGGCACTGACCTTCG 3045  
DB 3152 TGAAGCAAAAGAACAGTGTCTGACGAGACTGGGTTGAAGGCACTGACCTTCG 3211  
QY 3050 ATTGCTGTAGGACACTCTCTTGGCCCTTCTGCAACCGGATCTGACCACTTAAAC 3109  
DB 3212 ATTGCTGTAGGACACTCTCTTGGCCCTTCTGCAACCGGATCTGACCACTTAAAC 3271  
QY 3110 TGGTGCAAAATTAATCTTCACTCCCAAGGAATGAAAGCTGTGCTGCTTGGCTGTC 3169  
DB 3272 TGGTGCAAAATTAATCTTCACTCCCAAGGAATGAAAGCTGTGCTGCTTGGCTGTC 3331  
QY 3170 CCAAGCTTAACTTACAGATTAATTTGGCTGTGAAATGAGAGTAACTTGTGCAAAATTAAGA 3229  
DB 3332 CCAAGCTTAACTTACAGATTAATTTGGCTGTGAAATGAGAGTAACTTGTGCAAAATTAAGA 3391  
QY 3320 AGCTCTGAGAGAGTGCAGTACTCAAGCCCGAGTCTGTAATTGACGGTAGTGGCATT 3289  
DB 3392 AGCTCTGAGAGAGTGCAGTACTCAAGCCCGAGTCTGTAATTGACGGTAGTGGCATT 3451  
QY 3290 CTTTGTATGATGATGACCG 3308  
DB 3452 CTTTGTATGATGATGACCG 3470

RESULT 5  
US-10-216-645-3

/ Sequence 3, Application US/10216645  
/ Publication No. US20030125282A1  
/ GENERAL INFORMATION:  
/ APPLICANT: WEISS, BERTRAM  
/ APPLICANT: LESSL, MONIKA  
/ APPLICANT: PETERS-KOTTIG, MICHAEL  
/ APPLICANT: BECKMANN, GEORG  
/ TITLE OF INVENTION: HUMAN MATER PROTEINS  
/ FILE REFERENCE: SCH-1910  
/ CURRENT APPLICATION NUMBER: US/10/216,645  
/ CURRENT FILING DATE: 2003-01-21  
/ PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
/ NUMBER OF SEQ ID NOS: 11  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 3  
/ LENGTH: 3830  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-10-216-645-3

Query Match 70.9%; Score 2862; DB 15; Length 3830;  
Beet Local Similarity 99.9%; Pred. No. 0;  
Matches 2962; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 345 AGAATTTTCAAGCTATGAGCAAGAGTGCACAGAGAGAGAGAGAGAGCAAGG 404  
DB 450 AAAAAATTTCAAGCTATGAGCAAGAGTGCACAGAGAGAGAGAGAGAGCAAGG 509  
QY 405 ACATGAGGATGACATGAGGACTACAGAGTGCAGTGAATGCAAAATTCGCTGAGAGGA 464  
DB 510 ACATGAGGATGACATGAGGACTACAGAGTGCAGTGAATGCAAAATTCGCTGAGAGGA 569  
QY 465 GATATGATGCTGATGTTTGAAGCACTGCTGCTGATGCTGCGGAAATGCAAACTTGGC 524  
DB 570 GATATGATGCTGATGTTTGAAGCACTGCTGCTGATGCTGCGGAAATGCAAACTTGGC 629  
QY 525 TGGTCTTTTGAATGACAGCGGCTTCCGCGCTGCAAGGATGCTTCTGCAAGGAA 584  
DB 630 TGGTCTTTTGAATGACAGCGGCTTCCGCGCTGCAAGGATGCTTCTGCAAGGAA 689  
QY 585 GTTCAAGATTTGGAATCGGCTCTAGCCAGAGATCGTGTGCTGCGGCGCAAGTGG 644  
DB 690 GTTCAAGATTTGGAATCGGCTCTAGCCAGAGATCGTGTGCTGCGGCGCAAGTGG 749

QY 645 ACTTACAGAGGAAATGTTCTCTAGCTCTTCTCTCCCGGTTAGAGATGACAGCGAA 704  
DB 750 ACTTACAGAGGAAATGTTCTCTAGCTCTTCTCTCCCGGTTAGAGATGACAGCGAA 809  
QY 705 GAAAGAGAGAGTGTCAAGAGTTCATCTCCAGAGAGTGGCCAGATCCAGGCTCCGGT 764  
DB 810 GAAAGAGAGAGTGTCAAGAGTTCATCTCCAGAGAGTGGCCAGATCCAGGCTCCGGT 869  
QY 765 GACGAGATCATGTCTCCAGCAAGAAAGGCTGTTTATCATATGACGGTTTGCATGACCT 824  
DB 870 GACGAGATCATGTCTCCAGCAAGAAAGGCTGTTTATCATATGACGGTTTGCATGACCT 929  
QY 825 GGGCTCTGCTTCAACATGACAAAGCTGTGCAAAAGTGGGCTGAGAGACCTCC 884  
DB 930 GGGCTCTGCTTCAACATGACAAAGCTGTGCAAAAGTGGGCTGAGAGACCTCC 989  
QY 885 GTTCAACCTTCAACAGCACTGTGAGAGAGTCTGCTCTCCCTGAGTCTTCTGATGCT 944  
DB 990 GTTCAACCTTCAACAGCACTGTGAGAGAGTCTGCTCTCCCTGAGTCTTCTGATGCT 1049  
QY 945 CACGCTCAGAGAGTGGGCAAGAGAGCTCAAGTCAAGGTCGTGTCTCCCGTTACT 1004  
DB 1050 CACGCTCAGAGAGTGGGCAAGAGAGCTCAAGTCAAGGTCGTGTCTCCCGTTACT 1109  
QY 1005 GTTACTTGAAGAAATCTCCGGGAAACAAGAAATCACTGCTCTTGAAGCCGGATGG 1064  
DB 1110 GTTACTTGAAGAAATCTCCGGGAAACAAGAAATCACTGCTCTTGAAGCCGGATGG 1169  
QY 1065 TGAGCATCAGAAAGACACAAGGGTGTGCGCATCATGAAACAACGCTGAGCTGTGACCA 1124  
DB 1170 TGAGCATCAGAAAGACACAAGGGTGTGCGCATCATGAAACAACGCTGAGCTGTGACCA 1229  
QY 1125 GTGCCAGTCCCGCGCTGTGCTCTCATCTGCTGCTGCTGACGCTGACAGCTGTGT 1184  
DB 1230 GTGCCAGTCCCGCGCTGTGCTCTCATCTGCTGCTGCTGACGCTGACAGCTGTGT 1289  
QY 1185 GGGGAGAGCGTGGCCCTTCAACCAAGCTCAAGGCTGACAGCGGCTTGTGT 1244  
DB 1290 GGGGAGAGCGTGGCCCTTCAACCAAGCTCAAGGCTGACAGCGGCTTGTGT 1349  
QY 1245 TCATGAGTCAACCCCTGAGGCGTGTGCTGCTGCTCATCTGAGAGAGAGTGT 1304  
DB 1350 TCATGAGTCAACCCCTGAGGCGTGTGCTGCTGCTCATCTGAGAGAGAGTGT 1409  
QY 1305 CTTGAAGGCTTCTGCGGTATGCTGTGAGAGAGTGTGAATGAGAAATGATGTTGA 1364  
DB 1410 CTTGAAGGCTTCTGCGGTATGCTGTGAGAGAGTGTGAATGAGAAATGATGTTGA 1469  
QY 1365 TGTGACGACCTCATGTGTTCAAGGACTGGGAGTCTGAGTCCGTCTGTGTTCAAT 1424  
DB 1470 CGGTGACGACCTCATGTGTTCAAGGACTGGGAGTCTGAGTCCGTCTGTGTTCAAT 1529  
QY 1425 GAAATCTTCTCCAGACAGCACTGAGAGAGTACTACCTTCTTCCACTGACT 1484  
DB 1530 GAAATCTTCTTCCAGACAGCACTGAGAGAGTACTACCTTCTTCCACTGACT 1589  
QY 1485 CCAGGACTTCTGTGCGGCTTGTACTAGTGTGAGAGGCTTGAAATGAGCCAGCTCT 1544  
DB 1590 CCAGGACTTCTGTGCGGCTTGTACTAGTGTGAGAGGCTTGAAATGAGCCAGCTCT 1649  
QY 1545 CTGCTCTTGTGAGTGAAGACAAAGAGTCCATGAGCTTAAACAGGCTTCCA 1604  
DB 1650 CTGCTCTTGTGAGTGAAGACAAAGAGTCCATGAGCTTAAACAGGCTTCCA 1709  
QY 1605 TATCAACTGCTTGAAGAGAGGCTTCTGTTGGCTGTGAGCGGAAGCTTAAGAG 1664  
DB 1710 TATCAACTGCTTGAAGAGAGGCTTCTGTTGGCTGTGAGCGGAAGCTTAAGAG 1769  
QY 1665 GCACTGAGAGTCTGTGCTGCTGTCCGTTCCCTGAGGAGTGAAGAGAGCTTCTGCA 1724  
DB 1770 GCACTGAGAGTCTGTGCTGCTGTCCGTTCCCTGAGGAGTGAAGAGAGCTTCTGCA 1829  
QY 1725 CTGGGTCTCTGTGTTGGGTCAAGAGCTTAATGCCAACCCAGAGAGACCTTGAAGCC 1784



Db 1830 CTGGGTCTCTGTTGGGTCAAGCATTAATGCCACCAAGGAGACCCCTGAGACG 1889  
Qy 1785 CTTCACATGTCCTTTTCAGAGCTCAAGACAAAGATTGTTGGCTTGGATTAACAGCTT 1844  
Db 1890 CTTCACATGTCCTTTTCAGAGCTCAAGACAAAGATTGTTGGCTTGGATTAACAGCTT 1949  
Qy 1845 CCAAGAAATGTCCTTCGATTAACCAAGACCTGGACTTGTAGCATTTCTCTGCT 1904  
Db 1950 CCAAGAAATGTCCTTCGATTAACCAAGACCTGGACTTGTAGCATTTCTCTGCT 2009  
Qy 1905 CCAGACATGTCCTTAATTTGCGGAAAATTCGAGTGTCAAGAGGATCTTCCAAAGAA 1964  
Db 2010 CCAGACATGTCCTTAATTTGCGGAAAATTCGAGTGTCAAGAGGATCTTCCAAAGAA 2069  
Qy 1965 TGAATCCCTGAGGAGATGTCCTGTCCTCTATAGATGCGGAGTAAGACCTCATTTGA 2024  
Db 2070 TGAATCCCTGAGGAGATGTCCTGTCCTCTATAGATGCGGAGTAAGACCTCATTTGA 2139  
Qy 2025 GAGAGAGTGGGAAATTTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGGA 2084  
Db 2130 GAGAGAGTGGGAAATTTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGGA 2189  
Qy 2085 CCTGGGAGAGCATCTGACAGAGCGGCGCATGAAAGACCTGTGTGCAAGCTGAGCA 2144  
Db 2190 CCTGGGAGAGCATCTGACAGAGCGGCGCATGAAAGACCTGTGTGCAAGCTGAGCA 2249  
Qy 2145 TCCCACTGCAAGATACAGACCTCTGATGTTTAAATGCAAGATTAACCTCTGTGTGCA 2204  
Db 2250 TCCCACTGCAAGATACAGACCTCTGATGTTTAAATGCAAGATTAACCTCTGTGTGCA 2309  
Qy 2205 GCACCTGTGAGAAATCGTCATGAGCCCAACCTTAAGATCCCTCAACTTGGAGGAGC 2264  
Db 2310 GCACCTGTGAGAAATCGTCATGAGCCCAACCTTAAGATCCCTCAACTTGGAGGAGC 2369  
Qy 2265 CCACCTGAGAGAGATGTAAGAGATGCGTGTGAAGCCTTAAACCAACCAAAATGTTT 2324  
Db 2370 CCACCTGAGAGAGATGTAAGAGATGCGTGTGAAGCCTTAAACCAACCAAAATGTTT 2429  
Qy 2335 GTTGAAGTCCTTGAAGGCTGAGATGTCGTGATGACCAATGCTTATCTGAAGATCTC 2384  
Db 2430 GTTGAAGTCCTTGAAGGCTGAGATGTCGTGATGACCAATGCTTATCTGAAGATCTC 2489  
Qy 2385 CCAATCTTGAAGCCTTCAAGCCTGAGATGTCGTGATGAGGAGGAGGAGGAGGAGG 2444  
Db 2490 CCAATCTTGAAGCCTTCAAGCCTGAGATGTCGTGATGAGGAGGAGGAGGAGGAGG 2549  
Qy 2445 AGACCAAGGAGTAATGCTCTCAAGTATGCTTGAAGTCTCCAGTGCCTGACAGAA 2504  
Db 2550 AGACCAAGGAGTAATGCTCTCAAGTATGCTTGAAGTCTCCAGTGCCTGACAGAA 2609  
Qy 2505 GCTGATATCTGAGAGATGTCGTGATGACCAAGGAGTCTGAGTCTGAGGCT 2564  
Db 2610 GCTGATATCTGAGAGATGTCGTGATGACCAAGGAGTCTGAGTCTGAGGCT 2669  
Qy 2565 GCTGAGCAACCGGAGCTTGAACACCTGTCCTTCAACCAAGGAGTCTGAGGAGGAGG 2624  
Db 2670 GCTGAGCAACCGGAGCTTGAACACCTGTCCTTCAACCAAGGAGTCTGAGGAGGAGG 2729  
Qy 2635 TGTAAATCTACTGTGATGATCCATGAGGCTTCCCACTGTAAGTCTGAGAGGCTGATCT 2684  
Db 2730 TGTAAATCTACTGTGATGATCCATGAGGCTTCCCACTGTAAGTCTGAGAGGCTGATCT 2789  
Qy 2685 GAATCAAGTCCACCTGAGCAAGGCTGAGTGTGTTTCTTGAACCTGAGCTTAAGGTA 2744  
Db 2790 GAATCAAGTCCACCTGAGCAAGGCTGAGTGTGTTTCTTGAACCTGAGCTTAAGGTA 2849  
Qy 2745 CTCAATGCTGAGCAAGGCTGAGCTTAAGTGAACCTGAGGAGCAATGAGGAGGAGCT 2804  
Db 2850 CTCAATGCTGAGCAAGGCTGAGCTTAAGTGAACCTGAGGAGCAATGAGGAGGAGCT 2909  
Qy 2805 TCTGTGCAAGGTCATGAGAGAACCATTTGTATCTCCAGAGCTGAGGTTGTAAAGTG 2864

Db 2910 TCTGTGCAAGGTCATGAGAGAACCATTTGTATCTCCAGAGCTGAGGTTGTAAAGTG 2969  
Qy 2865 TCATCTACCGCGCGGCTGTGAGATGTCCTGTGATCTCGAGAGCAGACACT 2924  
Db 2970 TCATCTACCGCGCGGCTGTGAGATGTCCTGTGATCTCGAGAGCAGACACT 3029  
Qy 2925 GAGAGCTGATCTCAAGCAATGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2984  
Db 3030 GAGAGCTGATCTCAAGCAATGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 3089  
Qy 2985 GGAAGCTGAGCAAGAAAGACGTCCTTGAAGAGCTGAGAGCTGAGAGCTGAGC 3044  
Db 3090 GGAAGCTGAGCAAGAAAGACGTCCTTGAAGAGCTGAGAGCTGAGAGCTGAGC 3149  
Qy 3045 TTCTGATGCTGAGAGCACTCTGCTGAGGCTTCTGCAACCGGAGTCAAGCT 3104  
Db 3150 TTCTGATGCTGAGAGCACTCTGCTGAGGCTTCTGCAACCGGAGTCAAGCT 3209  
Qy 3105 AAACCTGTCAGAAATTAATTCAGTCCCAAGAAATGATGAGCTGTTCGAGCTTTCG 3164  
Db 3210 AAACCTGTCAGAAATTAATTCAGTCCCAAGAAATGATGAGCTGTTCGAGCTTTCG 3269  
Qy 3165 CTGTCCACGCTTAACCTTAACGATTAATGGCTGTGAAATGCGATACCTGTGCAAT 3224  
Db 3270 CTGTCCACGCTTAACCTTAACGATTAATGGCTGTGAAATGCGATACCTGTGCAAT 3329  
Qy 3225 AAGGAGCTGTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3284  
Db 3330 AAGGAGCTGTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3389  
Qy 3285 GCATTCCTTTGATGATGATGACCG 3308  
Db 3390 GCATTCCTTTGATGATGATGACCG 3413

RESULT 6  
US-10-092-900A-347  
Sequence 347, Application US/10092900A  
Publication No. US20040043382A1  
GENERAL INFORMATION:  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Spytek, Kimberly A.  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Taupier Jr., Raymond J.  
APPLICANT: Pena, Carol E.A.  
APPLICANT: Li, Li  
APPLICANT: Zernusen, Bryan D.  
APPLICANT: Gusev, Vladimir Y.  
APPLICANT: Ji, Weizhen  
APPLICANT: Gorman, Linda  
APPLICANT: Miller, Charles E.  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Paturajan, Meera  
APPLICANT: Gangolli, Bsha A.  
APPLICANT: Vernet, Corine A.M.  
APPLICANT: Guo, Xiaojia Sasha  
APPLICANT: Tchernev, Velizar T.  
APPLICANT: Fernandez, Rima R.  
APPLICANT: Casman, Stacie J.  
APPLICANT: Malyankar, Uriel M.  
APPLICANT: Gerlach, Valerie  
APPLICANT: Liu, Yi  
APPLICANT: Anderson, David W.  
APPLICANT: Spaderma, Steven K.  
APPLICANT: Catterton, Elina  
APPLICANT: Leite, Mario W.  
APPLICANT: Zhong, Haihong  
APPLICANT: Alsbrook, John P.  
APPLICANT: Lepley, Denise M.  
APPLICANT: Rieger, Daniel K.  
APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US20040043382A1 Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C



CURRENT APPLICATION NUMBER: US/10/092,900A  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: USSN 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: USSN 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: USSN 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: USSN 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: USSN 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: USSN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USSN 60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See file Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 347  
LENGTH: 3226  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (69)..(3168)  
US-10-092-900A-347

Query Match 69.4%; Score 2801; DB 18; Length 3226;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2901; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 406 CATGAGGTGACACATGGGACTACAGAGTACCGTGAAGACCAATTTGCTGAGAGAG 465  
DB 249 CATGAGGTGACACATGGGACTACAGAGTACCGTGAAGACCAATTTGCTGAGAGAG 308  
QY 466 GATGACCTGCTAGTTTGAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525  
DB 309 GATGACCTGCTAGTTTGAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368  
QY 526 GGTGCTTTGATTCAGACCGGTGGGCTTCGCGCTCGACCGTGTCTGACCGGAAG 585  
DB 369 GGTGCTTTGATTCAGACCGGTGGGCTTCGCGCTCGACCGTGTCTGACCGGAAG 428  
QY 586 TCAGGAATTTGGGAATCGGCTCTAGCCAGAGAGATCGTGTGCTGCTGCGCGCAGCTGGA 645  
DB 429 TCAGGAATTTGGGAATCGGCTCTAGCCAGAGAGATCGTGTGCTGCTGCGCGCAGCTGGA 488  
QY 646 CTCTACAGGGAATGTTCTCTAGCTCTTCTCTCCCGTTAAGAGATGACAGCGAG 705  
DB 489 CTCTACAGGGAATGTTCTCTAGCTCTTCTCTCCCGTTAAGAGATGACAGCGAG 548  
QY 706 AAGAGAGCACTGTCACAGATTCCTCAAGAGTGGCCAGACTCCAGGCTCCGGTG 765  
DB 549 AAGAGAGCACTGTCACAGATTCCTCAAGAGTGGCCAGACTCCAGGCTCCGGTG 608  
QY 766 ACGAGATCATGTCCCGACCGAAGAGCTGTTGTTTATCATTTAGACGTTTGAATGACTG 825  
DB 609 ACGAGATCATGTCCCGACCGAAGAGCTGTTGTTTATCATTTAGACGTTTGAATGACTG 668  
QY 826 GGTCTGTCTCTCAACATGACACAAAGCTGTGCAAGAGCTGGGCTGAGAGAGCTTCGG 885  
DB 669 GGTCTGTCTCTCAACATGACACAAAGCTGTGCAAGAGCTGGGCTGAGAGAGCTTCGG 728  
QY 886 TTCAACCTCATAGCAGCTCTGCTGAGAGAGTCTGCTCTCTGAGTCTTCTGATCTGC 945  
DB 729 TTCAACCTCATAGCAGCTCTGCTGAGAGAGTCTGCTCTCTGAGTCTTCTGATCTGC 788

QY 946 ACCGTACAGAGCTGTGGGACAGAGAAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTG 1005  
DB 789 ACCGTACAGAGCTGTGGGACAGAGAAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTG 848  
QY 1006 TTAGTTAGAGAAATCTCCGGGGAACAAAGATTCACCTTGTCTCTTGAACCGCGGATTTG 1065  
DB 849 TTAGTTAGAGAAATCTCCGGGGAACAAAGATTCACCTTGTCTCTTGAACCGCGGATTTG 908  
QY 1066 GAGCATTCAGAAACAAAGGTTGCTGCTGATCATTAACAAACCTGAGCTGTGACACAG 1125  
DB 909 GAGCATTCAGAAACAAAGGTTGCTGCTGATCATTAACAAACCTGAGCTGTGACACAG 968  
QY 1126 TGCCAGGTGCGCGGTGCTCTCATCTGCGGCGCCGACGCTGACAGAGTGTG 1185  
DB 969 TGCCAGGTGCGCGGTGCTCTCATCTGCGGCGCCGACGCTGACAGAGTGTG 1028  
QY 1186 GGGAGAGCGTGTGCGCCCTTCAACCAACGCTCAACAGGCTGACGCGCTTTTGTGTT 1245  
DB 1029 GGGAGAGCGTGTGCGCCCTTCAACCAACGCTCAACAGGCTGACGCGCTTTTGTGTT 1088  
QY 1246 CATGAGCTCACCCCTCGAGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1305  
DB 1089 CATGAGCTCACCCCTCGAGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148  
QY 1306 CTGAAGCGCTTCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1365  
DB 1149 CTGAAGCGCTTCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208  
QY 1366 GGTGACGACCTCATGTTTCAAGAGCTGCGGAGTCTGAGCTGCTGCTGCTGCTGCTGCT 1425  
DB 1209 GGTGACGACCTCATGTTTCAAGAGCTGCGGAGTCTGAGCTGCTGCTGCTGCTGCTGCT 1268  
QY 1426 AACATCTTCTCCAGACAGGCACTGTGAGAGTACTACACTTCTTCCACTCATGCTC 1485  
DB 1269 AACATCTTCTCCAGACAGGCACTGTGAGAGTACTACACTTCTTCCACTCATGCTC 1328  
QY 1486 CAGACCTTGTGCGGCTTGTACTAGTGTTAAGGCGCTGGAATTCAGCCAGCTCTC 1545  
DB 1329 CAGACCTTGTGCGGCTTGTACTAGTGTTAAGGCGCTGGAATTCAGCCAGCTCTC 1388  
QY 1546 TGCCCTGTGAGTTGAAGACAAAGAGTCCATGAGCTTAAACAGGCGGCTTCAT 1605  
DB 1389 TGCCCTGTGAGTTGAAGACAAAGAGTCCATGAGCTTAAACAGGCGGCTTCAT 1448  
QY 1606 ATCCACTGCTTTGATGAAGCGTTTCTGTTGAGCTCGTGAAGAGATGAAGAG 1665  
DB 1449 ATCCACTGCTTTGATGAAGCGTTTCTGTTGAGCTCGTGAAGAGATGAAGAG 1508  
QY 1666 CCATCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1725  
DB 1509 CCATCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1568  
QY 1726 TGCGTCTCTGTTGGGTGAGCAGCTTAATGSCACACCCAGAGACACCTGGAAGGCC 1785  
DB 1569 TGCGTCTCTGTTGGGTGAGCAGCTTAATGSCACACCCAGAGACACCTGGAAGGCC 1628  
QY 1786 TTCCACTGTCTTTTGAAGTCAAGACAAAGATTTGCTGCTTGAACAGCTTC 1845  
DB 1529 TTCCACTGTCTTTTGAAGTCAAGACAAAGATTTGCTGCTTGAACAGCTTC 1688  
QY 1846 CAAAGATGTGAGCTTCCGATTAACAGAACTTGAATGATCTTCTTCTGCTC 1905  
DB 1689 CAAAGATGTGAGCTTCCGATTAACAGAACTTGAATGATCTTCTTCTGCTC 1748  
QY 1906 CAGCACTGTCCGATTTTGGGGAATAATCGGGTGAATGCAAGGGATCTTCCCAAGAT 1965  
DB 1749 CAGCACTGTCCGATTTTGGGGAATAATCGGGTGAATGCAAGGGATCTTCCCAAGAT 1808  
QY 1966 GAGTCCGCTGAGGATGCTGTGTGCTCTCTATGATGAGTGGGATTAAGACCTTCATTTAG 2025  
DB 1809 GAGTCCGCTGAGGATGCTGTGTGCTCTCTATGATGAGTGGGATTAAGACCTTCATTTAG 1868  
QY 2026 GAGCAGTGGGAAGATTTTGTGCTCATGTCTTGGACCAACCAACCACTGCGGAGCTGGAC 2085

Db 1869 GAGAGTGGGAAAGATTTCTGCTCCATGCTTGGCAACCCACCACTGCGGCACTGGAC 1928  
Qy 2086 CTGGGACAGACATCTCTGACAGAGCGGGCCATGAAAGCCCTGTGCGCAACTGAGCAT 2145  
Db 1929 CTGGGACAGACATCTCTGACAGAGCGGGCCATGAAAGCCCTGTGCGCAACTGAGCAT 1988  
Qy 2146 CCCACTGCAAGATACAGACCTGATGTTTGAATAATGACAGATTACCTGTGTGAG 2205  
Db 1989 CCCACTGCAAGATACAGACCTGATGTTTGAATAATGACAGATTACCTGTGTGAG 2048  
Qy 2206 CACCTTGGAGAAATGCTATGCGCAACCGTAACCTAAGATCCTCAACTTGGAGGAC 2265  
Db 2049 CACCTTGGAGAAATGCTATGCGCAACCGTAACCTAAGATCCTCAACTTGGAGGAC 2108  
Qy 2266 CACCTGAAAGAAAGGATGTAAGATGCGGTGTAAGCCCTTAAACCCCAAAATGTTG 2325  
Db 2109 CACCTGAAAGAAAGGATGTAAGATGCGGTGTAAGCCCTTAAACCCCAAAATGTTG 2168  
Qy 2326 TTGAGATCTTGAAGCTGATTTGCTGTGATTTGACCCATGCTGTACCTGAAGATCTC 2385  
Db 2169 TTGAGATCTTGAAGCTGATTTGCTGTGATTTGACCCATGCTGTGTACCTGAAGATCTC 2228  
Qy 2386 CAAATCTTACGACCTCCCAAGCTGTAATCTGTAGCTGCGAGAAACAAGGTGACA 2445  
Db 2229 CAAATCTTACGACCTCCCAAGCTGTAATCTGTAGCTGCGAGAAACAAGGTGACA 2288  
Qy 2446 GACCAAGGAAATGATGCTCTGATGATGCTGTAAGCTGCTGCGGCGCTGACAGAG 2505  
Db 2289 GACCAAGGAAATGATGCTCTGATGATGCTGTAAGCTGCTGCGGCGCTGACAGAG 2348  
Qy 2506 CTGATTAAGAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2565  
Db 2349 CTGATTAAGAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2408  
Qy 2566 GTGAGCAACCGGAGCTTGAACAACCTGTGCTTATCAACAACCTGTGAGGAGAGAGGT 2625  
Db 2409 GTGAGCAACCGGAGCTTGAACAACCTGTGCTTATCAACAACCTGTGAGGAGAGAGGT 2468  
Qy 2626 GTAAATCTACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2685  
Db 2469 GTAAATCTACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2528  
Qy 2686 AATGAGTCCACTGAGCAACCGGCTGTGTGTTTCTTGTGACCTTGTGAGGATGATG 2745  
Db 2529 AATGAGTCCACTGAGCAACCGGCTGTGTGTTTCTTGTGACCTTGTGAGGATGATG 2588  
Qy 2746 TCATGCTGAGCACTGAGCTTGTGATGAAACCTGTGAGAAACAATGCGGTGAAGCTT 2805  
Db 2589 TCATGCTGAGCACTGAGCTTGTGATGAAACCTGTGAGAAACAATGCGGTGAAGCTT 2648  
Qy 2806 CTGTGCGAGATGATGAGAAACAATGCTGTAATGCACTGCGAGACTGTGAGTGAAGT 2865  
Db 2649 CTGTGCGAGATGATGAGAAACAATGCTGTAATGCACTGCGAGACTGTGAGTGAAGT 2708  
Qy 2866 CATCTCAACCGGCGGTGTGATGAGTCTGTCTGTGATCTGTGAGGAGAGACACTG 2925  
Db 2709 CATCTCAACCGGCGGTGTGATGAGTCTGTCTGTGATCTGTGAGGAGAGACACTG 2768  
Qy 2926 AAGAGCTGATCTCAACGCAATGCTGTGAGTGAAGGCTGTGCGCTGTGCGAG 2985  
Db 2769 AAGAGCTGATCTCAACGCAATGCTGTGAGTGAAGGCTGTGCGCTGTGCGAG 2828  
Qy 2986 GGAATGAGCAAAAGAAAGATGTTCTGACAGACTGCGGTGAGAGGATGAGTGAAGT 3045  
Db 2829 GGAATGAGCAAAAGAAAGATGTTCTGACAGACTGCGGTGAGAGGATGAGTGAAGT 2888  
Qy 3046 TCTGATGCTGTGAGGACTCTCTTGTGCGCTTCTGCTGCAACCGGCAATGACAGTCTA 3105  
Db 2889 TCTGATGCTGTGAGGACTCTCTTGTGCGCTTCTGCTGCAACCGGCAATGACAGTCTA 2948  
Qy 3106 AACCTGTGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 3165

Db 2949 AACCTGTGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 3008  
Qy 3166 TGTCCAGCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3225  
Db 3009 TGTCCAGCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3068  
Qy 3226 AGAAGCTGTGAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3285  
Db 3069 AGAAGCTGTGAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3128  
Qy 3286 CATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3308  
Db 3129 CATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3151

RESULT 7  
US-10-860-761-3  
; Sequence 3, Application US/10860761  
; Publication No. US2004024875A1  
; GENERAL INFORMATION:  
; APPLICANT: MYETH  
; TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS  
; FILE REFERENCE: AM101318  
; CURRENT APPLICATION NUMBER: US/10/860,761  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 3885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3603)  
US-10-860-761-3

Query Match 67.2%; Score 2713; DB 20; Length 3885;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 290 AATTTTCAACGATGATGAAACAAGAGTGTGCTGACAGCAAGAGACAGAAACAAGAAA 349  
Db 566 AATTTTCAACGATGATGAAACAAGAGTGTGCTGACAGCAAGAGACAGAAACAAGAAA 625  
Qy 350 TTTTCAAGCTATGAAACAAGAGTGTGCTGACAGCAAGAGACAGAAACAAGAGCATG 409  
Db 626 TTTTCAAGCTATGAAACAAGAGTGTGCTGACAGCAAGAGACAGAAACAAGAGCATG 685  
Qy 410 GAGGTGACATGAGGACTTACAGAGTGTGATGATGATGATGATGATGATGATGATGATG 469  
Db 686 GAGGTGACATGAGGACTTACAGAGTGTGATGATGATGATGATGATGATGATGATGATG 745  
Qy 470 TACGTGTGATTTTGAAGAAACAAGTGTGCTGATGATGATGATGATGATGATGATGATG 529  
Db 746 TACGTGTGATTTTGAAGAAACAAGTGTGCTGATGATGATGATGATGATGATGATGATG 805  
Qy 530 CTTTGAATGACAGCGGTGAGGCTTCCGCGCTGCAAGGATGATGATGATGATGATGATGATG 589  
Db 806 CTTTGAATGACAGCGGTGAGGCTTCCGCGCTGCAAGGATGATGATGATGATGATGATGATG 865  
Qy 590 GAATTTGAAATGAGGCTTACAGAGATGATGATGATGATGATGATGATGATGATGATGATG 649  
Db 866 GAATTTGAAATGAGGCTTACAGAGATGATGATGATGATGATGATGATGATGATGATGATG 925  
Qy 650 ACCAGGAAATGTTCTCTTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 709  
Db 926 ACCAGGAAATGTTCTCTTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 985  
Qy 710 AGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 769  
Db 986 AGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1045  
Qy 770 AGATCATGTCCGACAGAAAGCTGTGTTTCAATGATGATGATGATGATGATGATGATGATG 829

Db	1046	AGATCATGTC	CCGACCA	GAAGAG	CGTGTGTT	ATCAT	TGACGG	TTTGCAT	TGAC	CTGG	CGT	1105									
Qy	830	CTGT	CCCTCA	CAATGA	CA	CAAA	AGCT	CTG	CAAA	GA	CTGG	CTGA	GAAG	CA	CGCT	CCGTT	CA	889			
Db	1106	CTGT	CTCA	CAATGA	CA	CAAA	AGCT	CTG	CAAA	GA	CTGG	CTGA	GAAG	CA	CGCT	CCGTT	CA	1165			
Qy	890	CCCT	CATA	CGA	GTCT	GTGA	GAAG	GTCT	GTCT	CCCT	GA	GTCT	CTT	CGTAT	CGT	CAC	CG	949			
Db	1166	CCCT	CAT	ACG	CAGT	CTG	CTGA	GAAG	GTCT	GTCT	CCCT	GA	GTCT	CTT	CGTAT	CGT	CAC	CG	1222		
Qy	950	TCAG	AGA	CGT	GGG	CA	CA	GAAG	GTCT	CAAGT	CGT	GTCT	CCCG	TTA	AC	CTGTT	AG	1008			
Db	1226	TCAG	AGA	CGT	GGG	CA	CA	GAAG	GTCT	CAAGT	CGT	GTCT	CCCG	TTA	AC	CTGTT	AG	1285			
Qy	1010	TTAG	AGAA	TC	TCGG	GGGA	CAAA	GAAT	CA	CTT	GTCT	CTT	GAGG	CGGG	AA	TTGG	TGAG	1065			
Db	1286	TTAG	AGAA	TC	TCGG	GGGA	CAAA	GAAT	CA	CTT	GTCT	CTT	GAGG	CGGG	AA	TTGG	TGAG	1345			
Qy	1070	ATCAGA	AGA	CA	CAAG	GGTTCG	TGCGAT	CAT	GA	CA	CCGT	GA	GTCT	CGA	CCAGT	CC	1129				
Db	1346	ATCAGA	AGA	CA	CAAG	GGTTCG	TGCGAT	CAT	GA	CA	CCGT	GA	GTCT	CGA	CCAGT	CC	1405				
Qy	1130	AGGT	CCCG	CCGT	GGG	CTCT	CTCAT	CTG	GTG	GGCC	CTG	CA	GTG	CA	GA	CGT	GTG	GGGG	1185		
Db	1406	AGGT	CCCG	CCGT	GGG	CTCT	CTCAT	CTG	GTG	GGCC	CTG	CA	GTG	CA	GA	CGT	GTG	GGGG	1465		
Qy	1190	AGAG	GTG	CGCC	CCCTT	CAAC	CAAA	GGCT	CA	CA	GGG	CTG	CA	CGG	CCCTT	TGT	GTAT	CAT	1245		
Db	1466	AGAG	GTG	CGCC	CCCTT	CAAC	CAAA	GGCT	CA	CA	GGG	CTG	CA	CGG	CCCTT	TGT	GTAT	CAT	1525		
Qy	1250	AGCT	CA	CCCT	CTGA	GGCG	GTG	TC	CGG	CCGT	CTCAT	CTG	AG	CAAA	GAAG	AT	GTG	CTGA	1309		
Db	1526	AGCT	CA	CCCT	CTGA	GGCG	GTG	TC	CGG	CCGT	CTCAT	CTG	AG	CAAA	GAAG	AT	GTG	CTGA	1585		
Qy	1310	AGCG	CTT	CTG	CCGTA	TGCGT	GTGA	GGAG	GTG	GAAT	AGAA	GT	CA	GT	TTT	GTAT	GTG	1365			
Db	1586	AGCG	CTT	CTG	CCGTA	TGCGT	GTGA	GGAG	GTG	GAAT	AGAA	GT	CA	GT	TTT	GTAT	GTG	1645			
Qy	1370	ACGA	CTCA	TGG	TTA	GAAG	CA	TCGG	GGG	AGT	CTGA	CT	CCG	GT	CTG	TTT	CA	ATGA	CA	1429	
Db	1646	ACGA	CTCA	TGG	TTA	GAAG	CA	TCGG	GGG	AGT	CTGA	CT	CCG	GT	CTG	TTT	CA	ATGA	CA	1705	
Qy	1430	TCCT	CT	CCCA	GAC	AGC	CACT	GTG	AG	AGT	ACTA	CA	CC	TT	CA	CT	CAGT	CT	CAAG	1489	
Db	1706	TCCT	CT	CCCA	GAC	AGC	CACT	GTG	AG	AGT	ACTA	CA	CC	TT	CA	CT	CAGT	CT	CAAG	1755	
Qy	1490	ACTT	CT	GTG	CCGCTT	GTAT	CTAC	GTGTT	AG	AGG	GCCT	GA	AA	TC	GA	GC	CA	GT	CT	GTCC	1549
Db	1766	ACTT	CT	GTG	CCGCTT	GTAT	CTAC	GTGTT	AG	AGG	GCCT	GA	AA	TC	GA	GC	CA	GT	CT	GTCC	1825
Qy	1550	CTCT	CT	GTAC	GTGA	GAAG	CAAA	GAAG	TC	CT	GA	AGCTT	AA	CA	GG	CA	AGCTT	CCAT	AT	CC	1609
Db	1826	CTCT	CT	GTAC	GTGA	GAAG	CAAA	GAAG	TC	CT	GA	AGCTT	AA	CA	GG	CA	AGCTT	CCAT	AT	CC	1885
Qy	1610	ACTG	CTT	TGAT	TGA	AGCGTTT	CTTGTT	TGG	CTGT	GA	GGAGA	GA	AGTA	GA	AGG	CGAC	1669				
Db	1886	ACTG	CTT	TGAT	TGA	AGCGTTT	CTTGTT	TGG	CTGT	GA	GGAGA	GA	AGTA	GA	AGG	CGAC	1945				
Qy	1670	TGAG	AGT	CTG	TGAG	CGTGT	CCCGTT	CCCT	CTG	GGGGT	GAAG	CA	GA	AGCTT	CTG	CA	CTGG	1729			
Db	1946	TGAG	AGT	CTG	TGAG	CGTGT	CCCGTT	CCCT	CTG	GGGGT	GAAG	CA	GA	AGCTT	CTG	CA	CTGG	2005			
Qy	1730	TCT	CTCT	GTGGT	TCAG	CA	AGCTT	AT	TGCA	CA	CCCC	CA	GA	GA	CA	CCCT	TCGA	CGCTT	CC	1789	
Db	2006	TCT	CTCT	GTGGT	TCAG	CA	AGCTT	AT	TGCA	CA	CCCC	CA	GA	GA	CA	CCCT	TCGA	CGCTT	CC	2065	
Qy	1790	ACTG	CTT	TCGA	GA	CTCA	GA	CAAA	GA	GT	TTGTT	CG	CT	GG	CA	TTTAA	CA	AGCTT	CCAA	1849	
Db	2066	ACTG	CTT	TCGA	GA	CTCA	GA	CAAA	GA	GT	TTGTT	CG	CT	GG	CA	TTTAA	CA	AGCTT	CCAA	2125	
Qy	1850	AA	GTGTG	CGTT	CCGAT	TAA	CCGA	AG	CT	TGA	AGT	CA	TT	CTT	CTT	CTG	CC	AG	1909		

Db	2126	AAGTGTGGCTTCCGATTAAACGAAACCTGGACCTTGATAGCATCTTCTGCTCCAGC	2185
Qy	1910	ACTGTCCGTAATTTGGGAAATTTGGGGTGATGTCAAAAGGATCTTCCAGATGAGT	1969
Db	2166	ACTGTCCGTAATTTGGGAAATTTGGGGTGATGTCAAAAGGATCTTCCAGATGAGT	2245
Qy	1970	CCGCTGAAGCATGTCTGTGTGATCCCTCTATGATGCGGGATTAAGACCTCTCATTTAGAGC	2029
Db	2246	CCGCTGAAGCATGTCTGTGTGATCCCTCTATGATGCGGGATTAAGACCTCTCATTTAGAGC	2305
Qy	2030	AGTGGGAAGATTTCTGTCTCCATNGCTGGCAACCCACACATCGCGGAGGTGGACCTGG	2089
Db	2306	AGTGGGAAGATTTCTGTCTCCATNGCTGGCAACCCACACATCGCGGAGGTGGACCTGG	2365
Qy	2090	GCAGAGCATCCTGACACAGAGCGGACCATGAAGAACCCTGTGTGCGCAAGCTGAGCATCCCA	2149
Db	2366	GCAGAGCATCCTGACACAGAGCGGACCATGAAGAACCCTGTGTGCGCAAGGTGAGCATCCCA	2425
Qy	2150	CCTGCAAGATACAGACCTGTGATGTTTAAAGATGACAGATTAACCCCTGGTGTGACACCC	2209
Db	2426	CCTGCAAGATACAGACCTGTGATGTTTAAAGATGACAGATTAACCCCTGGTGTGACACCC	2485
Qy	2210	TCTGGAGATGTGTCATGGCCCAACCGTAACTTAAATCCTCTCAACTTGGGAGGACCCAC	2265
Db	2486	TCTGGAGATGTGTCATGGCCCAACCGTAACTTAAATCCTCTCAACTTGGGAGGACCCAC	2545
Qy	2270	TGAAGGAAGAGATGTAAAGATGGGCGGTGAAGCCTTAAACACCCAAATGTTGTGTGG	2329
Db	2546	TGAAGGAAGAGATGTAAAGATGGGCGGTGAAGCCTTAAACACCCAAATGTTGTGTGG	2605
Qy	2330	AGTCTTTGAGGCTGGATTTGCTGTGATTTGACCCATGCTGTTTACTGTGAAGATCTCCCAA	2389
Db	2606	AGTCTTTGAGGCTGGATTTGCTGTGATTTGACCCATGCTGTTTACTGTGAAGATCTCCCAA	2665
Qy	2390	TCTTTAGCATCTCCCCAGCCTGAATCTCTGAGCTGTGCAGGAACAAGTGCACAGCC	2449
Db	2666	TCTTTAGCATCTCCCCAGCCTGAATCTCTGAGCTGTGCAGGAACAAGTGCACAGCC	2725
Qy	2450	AGGAGATTAAGCCTCTGAGTGAAGTCCTTGAAGATCTCCAGTGGGCCCTGCAGAAAGCTGA	2509
Db	2726	AGGAGATTAAGCCTCTGAGTGAAGTCCTTGAAGATCTCCAGTGGGCCCTGCAGAAAGCTGA	2785
Qy	2510	TACTGAGAGCATGTGGCATCAACAGCCAGGTTGCCAGATGTGGCCTCAGCCCTCTGCA	2569
Db	2786	TACTGAGAGCATGTGGCATCAACAGCCAGGTTGCCAGATGTGGCCTCAGCCCTCTGCA	2845
Qy	2570	GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAAGCTTGGGGAACGAAGGTGTA	2629
Db	2846	GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAAGCTTGGGGAACGAAGGTGTA	2905
Qy	2630	ATCTAATGTGTGCATTCATGAGGCTTCCCATCTTAATCTGCAAGGCTGATGTCTGAATC	2689
Db	2906	ATCTAATGTGTGCATTCATGAGGCTTCCCATCTTAATCTGCAAGGCTGATGTCTGAATC	2965
Qy	2690	AGTGCACCTTGAACAAGGCTGGCTGTGATTTCTTGGACATTGACGCTTAATGGGTAACTCAT	2749
Db	2966	AGTGCACCTTGAACAAGGCTGGCTGTGATTTCTTGGACATTGACGCTTAATGGGTAACTCAT	3025
Qy	2750	GAGTGAAGCATCTAGGCTTAAAGCATGAACCCCTGTGAAGAACAATGAGCGTGAAGCTTCTGT	2809
Db	3026	GAGTGAAGCATCTAGGCTTAAAGCATGAACCCCTGTGAAGAACAATGAGCGTGAAGCTTCTGT	3085
Qy	2810	GCGAGGTCAATGAGGAACAATCTTGTGTCACTCCAGAACCTGGAATTTGTAAAGTGTCAATC	2869
Db	3086	GCGAGGTCAATGAGGAACAATCTTGTGTCACTCCAGAACCTGGAATTTGTAAAGTGTCAATC	3145
Qy	2870	TCAACCGCGCGTGTGATGAGTCTGTCTGTGTGATCTCGAAGAGACAGACACTTGAAGA	2929
Db	3146	TCAACCGCGCGTGTGATGAGTCTGTCTGTGTGATCTCGAAGAGACAGACACTTGAAGA	3205
Qy	2930	GCTGTGATCTCAAGCAATATGCTCTGTGTGACGGTGTGGAGTTTGCTGCGCTGTGCGAGGAC	2989
Db	3206	GCTGTGATCTCAAGCAATATGCTCTGTGTGACGGTGTGGAGTTTGCTGCGCTGTGCGAGGAC	3265

Qy 2990 TGAAGCAAAAGNACAGTGTTCGACGAGACTCGGTTGAAGCATGTGACTGACTTCTG 3049  
Db 3266 TGAAGCAAAAGNACAGTGTTCGACGAGACTCGGTTGAAGCATGTGACTGACTTCTG 3325  
Qy 3050 ATTGCTGTGAGGACCTCTCTTGGCCCTTTCCTGACCCGCGATCTGACAGTCTAAAC 3109  
Db 3326 ATTGCTGTGAGGACCTCTCTTGGCCCTTTCCTGACCCGCGATCTGACAGTCTAAAC 3385  
Qy 3110 TGGTGCACAAATTAATCTTCAAGTCCCAAGAAATGAATGAATGATGTTGGCCCTTTCCTG 3169  
Db 3386 TGGTGCACAAATTAATCTTCAAGTCCCAAGAAATGAATGAATGATGTTGGCCCTTTCCTG 3445  
Qy 3170 CCACGCTTAACCTTACAGATTAATTTGGGCTGTGAAATGCGAGTACCTCTGCAAAATGAAGA 3229  
Db 3446 CCACGCTTAACCTTACAGATTAATTTGGGCTGTGAAATGCGAGTACCTCTGCAAAATGAAGA 3505  
Qy 3230 AGCTGCTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGTAATTGACGGTATGTTGGCATT 3289  
Db 3506 AGCTGCTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGTAATTGACGGTATGTTGGCATT 3565  
Qy 3290 CTTTGTGATGAAGATGACCG 3308  
Db 3566 CTTTGTGATGAAGATGACCG 3584

RESULT 8  
US-10-399-443-23  
Sequence 23, Application US/10399443  
Publication No. US20040028669A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as Represented by the  
APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
APPLICANT: Nelson, Lawrence M.  
APPLICANT: Tong, Zhi-Bin  
APPLICANT: Nelson, Lawrence  
APPLICANT: Zhi-Bin, Tong  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64785  
CURRENT APPLICATION NUMBER: US/10/399,443  
CURRENT FILING DATE: 2003-04-16  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 23  
LENGTH: 3900  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(3600)  
OTHER INFORMATION:  
US-10-399-443-23

Query Match 67.2%; Score 2713; DB 17; Length 3900;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 290 AAATTTCAGAGCTATGGAACAAGAGGTGCCAGACGACGACAGACAGACAGAAACAAGA 349  
Db 566 AAATTTCAGAGCTATGGAACAAGAGGTGCCAGACGACGACAGACAGACAGAAACAAGA 625  
Qy 350 TTTTCAAGCTATGGAACAAGAGGTGCCAGACGACGACAGACAGACAGAAACAAGGACATG 409  
Db 626 TTTTCAAGCTATGGAACAAGAGGTGCCAGACGACGACAGACAGACAGAAACAAGGACATG 685  
Qy 410 GAGTGACACATGGGACTTACAAAGTCAAGTGTATGATGATGATGATGATGATGATGATG 469  
Db 686 GAGTGACACATGGGACTTACAAAGTCAAGTGTATGATGATGATGATGATGATGATGATG 745

Qy 470 TACGTGTAGTTTGAAGAACATGCTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529  
Db 746 TACGTGTAGTTTGAAGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805  
Qy 530 CTTTGTATGACACCGGTGGGAGCTTCCGGCTCTGACAGTGTGTTCTGACAGGAAATGACG 589  
Db 806 CTTTGTATGACACCGGTGGGAGCTTCCGGCTCTGACAGTGTGTTCTGACAGGAAATGACG 865  
Qy 590 GAATTGGAAATATGGGCTCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 649  
Db 866 GAATTGGAAATATGGGCTCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 925  
Qy 650 ACCAGGAATGTTCTCTAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 709  
Db 926 ACCAGGAATGTTCTCTAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 985  
Qy 710 AGAGCAGTTCACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACCG 769  
Db 986 AGAGCAGTTCACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACCG 1045  
Qy 770 AGATCATGTCCCGACCAAGAGCTGTGTATCATTTGACGGTGTGATGATGATGATGATGAT 829  
Db 1046 AGATCATGTCCCGACCAAGAGCTGTGTATCATTTGACGGTGTGATGATGATGATGATGAT 1105  
Qy 830 CTGTCTCAACATGACACAAAGCTCTGCAAGACTGAGGCTGAGAACAGGCTCTCTCTCA 889  
Db 1106 CTGTCTCAACATGACACAAAGCTCTGCAAGACTGAGGCTGAGAACAGGCTCTCTCTCA 1165  
Qy 890 CCTCTATACGAGTCTGTGAGAGAGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 949  
Db 1166 CCTCTATACGAGTCTGTGAGAGAGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1225  
Qy 950 TCAAGACGTGGGACAGAGAGTCAAGTCAAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 1009  
Db 1226 TCAAGACGTGGGACAGAGAGTCAAGTCAAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 1285  
Qy 1010 TTAAGAGATCTCCGGGAAACAAGATCACTTGTCTCTTGAAGCGGAGTTGGTAGC 1069  
Db 1286 TTAAGAGATCTCCGGGAAACAAGATCACTTGTCTCTTGAAGCGGAGTTGGTAGC 1345  
Qy 1070 ATCAGAAACACAGAGGTGTGTCATCAATCAACACCTGAGTCTCTGACAGTGTGCC 1129  
Db 1346 ATCAGAAACACAGAGGTGTGTCATCAATCAACACCTGAGTCTCTGACAGTGTGCC 1405  
Qy 1130 AGGTGCGCGCGTGGGCTCTCTCAATGCGGCGCGTGCAGCTGCAAGAGTGTGGGGG 1189  
Db 1406 AGGTGCGCGCGTGGGCTCTCTCAATGCGGCGCGTGCAGCTGCAAGAGTGTGGGGG 1465  
Qy 1190 AGAGCTGTGCGCCCTTCAACCAACGCTCACAGGCTGCAAGCGCTTGTGTGTATC 1249  
Db 1466 AGAGCTGTGCGCCCTTCAACCAACGCTCACAGGCTGCAAGCGCTTGTGTGTATC 1525  
Qy 1250 AGCTCACCTCTGAGGCGTGTGCGGCGCTGTCTCAATCTGAGAAAGAGTGTCTCTGA 1309  
Db 1526 AGCTCACCTCTGAGGCGTGTGCGGCGCTGTCTCAATCTGAGAAAGAGTGTCTCTGA 1585  
Qy 1310 AGCGCTTCTGCGGATGCTGTGAGGAGTGTGAGGATGAGGATGAGGATGAGGATGAGG 1369  
Db 1586 AGCGCTTCTGCGGATGCTGTGAGGAGTGTGAGGATGAGGATGAGGATGAGGATGAGG 1645  
Qy 1370 ACAGCTCATATGTTCAAGGACTCGGGAGTCTGAGTCCGCTCTCTCTCTCTCTCTCTCT 1429  
Db 1646 ACAGCTCATATGTTCAAGGACTCGGGAGTCTGAGTCCGCTCTCTCTCTCTCTCTCTCT 1705  
Qy 1430 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACCTTCTCTCACTCATGTTCCAGG 1489  
Db 1706 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACCTTCTCTCACTCATGTTCCAGG 1765  
Qy 1490 ACTTCTGTGCGGCTTGTATCTAGTGTGAGGCGCTGAGAAATGAGACAGCTCTCTGCC 1549  
Db 1766 ACTTCTGTGCGGCTTGTATCTAGTGTGAGGCGCTGAGAAATGAGACAGCTCTCTGCC 1825

1550 CTCCTGATCTGAGAGCAAGAGGCTCATGAGCTTAAACAGGAGGCTTCCATATCC 1609  
1826 CTCCTGATCTGAGAGCAAGAGGCTCATGAGCTTAAACAGGAGGCTTCCATATCC 1885  
1610 ACTGCTTTGAGAGGCGTTTCTTGTGTCCTGTCGAGCGAAGAGCTAAGAGGCGAC 1669  
1886 ACTGCTTTGAGAGGCGTTTCTTGTGTCCTGTCGAGCGAAGAGCTAAGAGGCGAC 1945  
1670 TGGAGGCTCTGTCGAGGCTTCCGTTCCCTTGGGGGTGAAGAGAGCTTCTGCACTGG 1729  
1946 TGGAGGCTCTGTCGAGGCTTCCGTTCCCTTGGGGGTGAAGAGAGCTTCTGCACTGG 2005  
1730 TCTCTCTGTCGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAG 1789  
2006 TCTCTCTGTCGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAG 2065  
1790 ACTGCTTTGAGAGGCGTTTCTTGTGTCCTGTCGAGCGAAGAGCTAAGAGGCGAC 1849  
2066 ACTGCTTTGAGAGGCGTTTCTTGTGTCCTGTCGAGCGAAGAGCTAAGAGGCGAC 2125  
1850 AAGTGTGCTTCCGATTAACGAGAACCTGAGCTTGAATGAGCTTCTTCTGCTCCAGC 1909  
2126 AAGTGTGCTTCCGATTAACGAGAACCTGAGCTTGAATGAGCTTCTTCTGCTCCAGC 2185  
1910 ACTGCTTTGAGAGGCGTTTCTTGTGTCCTGTCGAGCGAAGAGCTAAGAGGCGAC 1969  
2186 ACTGCTTTGAGAGGCGTTTCTTGTGTCCTGTCGAGCGAAGAGCTAAGAGGCGAC 2245  
1970 CCGCTGAGGAGCTGCTGTCGTCCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2029  
2246 CCGCTGAGGAGCTGCTGTCGTCCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2305  
2030 AAGTGTGAGATTTCTGCTCACTGCTTGGAGCCCACTGAGGAGGAGGAGGAGGAGG 2089  
2306 AAGTGTGAGATTTCTGCTCACTGCTTGGAGCCCACTGAGGAGGAGGAGGAGGAGG 2365  
2090 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2149  
2366 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2425  
2150 CCGCTGAGGAGCTGCTGTCGTCCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2209  
2426 CCGCTGAGGAGCTGCTGTCGTCCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2485  
2210 TCTGAGAGATGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2269  
2486 TCTGAGAGATGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2545  
2270 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2329  
2546 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2605  
2330 AGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2389  
2606 AGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2665  
2390 TCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2449  
2666 TCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2725  
2450 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2509  
2726 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2785  
2510 TACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2569  
2786 TACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2845  
2570 GCAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2629  
2846 GCAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2905  
2630 ATCTACTGTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2689

2906 ATCTACTGTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2965  
2690 AGTGTACCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2749  
2966 AGTGTACCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3025  
2750 GGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2809  
3026 GGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3085  
2810 GCGAGGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2869  
3086 GCGAGGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3145  
2870 TCACGCGCGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2929  
3146 TCACGCGCGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3205  
2930 GCTGTGATCTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2989  
3206 GCTGTGATCTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3265  
2990 TGAAGCAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3049  
3266 TGAAGCAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3325  
3050 ATTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3109  
3326 ATTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3385  
3110 TGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3169  
3386 TGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3445  
3170 CCAAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 3229  
3446 CCAAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 3505  
3220 AGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3289  
3506 AGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3565  
3290 CTTTGTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3308  
3566 CTTTGTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3584

RESULT 9  
US-10-677-943-23  
Sequence 23, Application US/10677943  
Publication No. US20040072297A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as  
APPLICANT: represented by the Secretary of the Department of Health and  
APPLICANT: Human Services  
APPLICANT: Nelson, Lawrence  
APPLICANT: Tong, Zhi-Bin  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64790  
CURRENT FILING DATE: US/10/677,943  
CURRENT FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US02/09776  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: Patent version 3.2  
SEQ ID NO 23  
LENGTH: 3900  
TYPE: DNA







2330 AGCTTTGAGGCTGATGCTGATGATGACCAATGCTGTTACCTGAAGATCTCCCAA 2389  
2606 AGCTTTGAGGCTGATGCTGATGATGACCAATGCTGTTACCTGAAGATCTCCCAA 2665  
2390 TCCCTTACGACCTCCCGACGCTGAAATCTCTGAGCCTGACAGAAAACAAGTGAAGACC 2449  
2666 TCCCTTACGACCTCCCGACGCTGAAATCTCTGAGCCTGACAGAAAACAAGTGAAGACC 2725  
2450 AGGAGGTAATGCTCTCTGATGATGCTGATGAGAGTCTCCAGAGTCCGCTGCAAGAGTCA 2509  
2726 AGGAGGTAATGCTCTCTGATGATGCTGATGAGAGTCTCCAGAGTCCGCTGCAAGAGTCA 2785  
2510 TACTGAGGACTGTGACATCAAGCAAGGTTGACAGAGTCCGCTGCAAGAGTCA 2569  
2786 TACTGAGGACTGTGACATCAAGCAAGGTTGACAGAGTCCGCTGCAAGAGTCA 2845  
2570 GCAACCGGAGCTTGACACACCTGTGCTTACCAACAACAGCCTGAGGAAACGAAGTGTAA 2629  
2846 GCAACCGGAGCTTGACACACCTGTGCTTACCAACAACAGCCTGAGGAAACGAAGTGTAA 2905  
2630 ATCTACTGTGTGATTCATGAGAGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTGAATC 2689  
2906 ATCTACTGTGTGATTCATGAGAGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTGAATC 2965  
2690 AGTGCACCTGAGACAGGCTGAGTGTGCTTCTGCACTTGCGCTTATGGTAACTCAT 2749  
2966 AGTGCACCTGAGACAGGCTGAGTGTGCTTCTGCACTTGCGCTTATGGTAACTCAT 3025  
2750 GGCCTGACGACCTGAGCCTTACATGAAACCTGTGGAAGACAATGAGCTGAACTTCTGT 2809  
3026 GGCCTGACGACCTGAGCCTTACATGAAACCTGTGGAAGACAATGAGCTGAACTTCTGT 3085  
2810 GCGAGGTCTAGAGAAACCATCTTGTCTATCTCCAGAGCCTGAGATTGTAAAGTGTATC 2869  
3086 GCGAGGTCTAGAGAAACCATCTTGTCTATCTCCAGAGCCTGAGATTGTAAAGTGTATC 3145  
2870 TCAACGCGCGGTGCTGTGAGTCTGTCTGTGTATCTCCAGAGCCTGAGATTGTAAAGTGTATC 2929  
3146 TCAACGCGCGGTGCTGTGAGTCTGTCTGTGTATCTCCAGAGCCTGAGATTGTAAAGTGTATC 3205  
2930 GGCCTGATCTACGACCAATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2989  
3206 GGCCTGATCTACGACCAATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3265  
2990 TGAAGCAAAAAGAAAGAGTGTCTGAGAGAGCTCGGTTGAAAGCATGTGAGTGTGAGTGTGAG 3049  
3266 TGAAGCAAAAAGAAAGAGTGTCTGAGAGAGCTCGGTTGAAAGCATGTGAGTGTGAGTGTGAG 3325  
3050 ATTGCTGTGAGGCACTCTCTGAGCCTTCTCTGCAACCGGATCTGACCAATCTAAACC 3109  
3326 ATTGCTGTGAGGCACTCTCTGAGCCTTCTCTGCAACCGGATCTGACCAATCTAAACC 3385  
3110 TGGTGCAGAAATTAATCTCACTGCTCCAAAGAGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 3169  
3386 TGGTGCAGAAATTAATCTCACTGCTCCAAAGAGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 3445  
3170 CCAAGCTTAACTTACAGATTAATTTGGCTGTGAGAAATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 3229  
3446 CCAAGCTTAACTTACAGATTAATTTGGCTGTGAGAAATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 3505  
3230 AGCTGTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3289  
3506 AGCTGTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3565  
3290 CTTTGTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3308  
3566 CTTTGTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3584

RESULT 10  
US-10-399-443-1  
; Sequence 1, Application US/10399443  
; Publication No. US20040028669A1

GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-399-443-1  
Query Match 27.3%; Score 1102; DB 17; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
757 GCTCCGAGAGGAGATCATCTCCGACCAAGAGGCTGTGTTTATCATGAGAGGTTTC 816  
4 GCTCCGAGAGGAGATCATCTCCGACCAAGAGGCTGTGTTTATCATGAGAGGTTTC 63  
817 GATGACTGAGGCTCTGTCTCTCAACATGACCAAGAGCTCTGCAAGAGCTGAGAG 876  
64 GATGACTGAGGCTCTGTCTCTCAACATGACCAAGAGCTCTGCAAGAGCTGAGAG 123  
877 CAGCTCCGTTCAACCTCTATACGAGTGTGTGAGAGAGTCTCTCTGAGAGTCTTC 936  
124 CAGCTCCGTTCAACCTCTATACGAGTGTGTGAGAGAGTCTCTCTGAGAGTCTTC 183  
937 CTGATGTCACCGTACAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 996  
184 CTGATGTCACCGTACAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 243  
997 GCTTACCTGTATTAGTGAAGAAATCTCCGAGGAAACAAAGATTCATTGCTCTTGAAGGC 1056  
244 GCTTACCTGTATTAGTGAAGAAATCTCCGAGGAAACAAAGATTCATTGCTCTTGAAGGC 303  
1057 GGGATGTGTGAGATCAGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1116  
304 GGGATGTGTGAGATCAGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 363  
1117 CTGACAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1176  
364 CTGACAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 423  
1177 GAGCTGTGTGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1236  
424 GAGCTGTGTGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 483  
1237 TTTGTGTTTATCAGCTCAACCTCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1296  
484 TTTGTGTTTATCAGCTCAACCTCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 543  
1297 AGAGTGTCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1356  
544 AGAGTGTCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 603  
1357 GTGTTTGAATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1416  
604 GTGTTTGAATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 663  
1417 TTTCAATGAACATCTTCTCTCCAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1476

Db 664 TTTCACATGAACATCTTCTCCAGACAGCACTGTAGAGATCTAACAACCTTCTTCCAC 723  
Qy 1477 CTCACCTTCACAGAGATCTCTGTGCCCCCTTGTACTACGTGTAGAGAGGCTTGAATTCAG 1536  
Db 724 CTCACCTTCACAGAGATCTCTGTGCCCCCTTGTACTACGTGTAGAGAGGCTTGAATTCAG 783  
Qy 1537 CCAGCTCTCTGCGCTCTGTACGTGTAGAGACAAAGAGGTCATGAGCTTAAACAGCA 1596  
Db 784 CCAGCTCTCTGCGCTCTGTACGTGTAGAGACAAAGAGGTCATGAGCTTAAACAGCA 843  
Qy 1597 GGCCTTCATATCCACTCGCTTGTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGAC 1656  
Db 844 GGCCTTCATATCCACTCGCTTGTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGAC 903  
Qy 1657 GTAAGAGAGCCCACTGAGAGTCTGTGAGGCTCTCCCGTCCCTGAGGGGTGAAGCAGAG 1716  
Db 904 GTAAGAGAGCCCACTGAGAGTCTGTGAGGCTCTCCCGTCCCTGAGGGGTGAAGCAGAG 963  
Qy 1717 CTTCGACACTGAGGCTCTGTGTTGGGTGAGCAGCCTTAATGCCACCCAGAGACACC 1776  
Db 964 CTTCGACACTGAGGCTCTGTGTTGGGTGAGCAGCCTTAATGCCACCCAGAGACACC 1023  
Qy 1777 CTGACGCGCTTCACACTGTCTTTTTCAGACTCAAGCAAAAGATTGTCGCTTGACATTA 1836  
Db 1024 CTGACGCGCTTCACACTGTCTTTTTCAGACTCAAGCAAAAGATTGTCGCTTGACATTA 1083  
Qy 1837 AACACCTTCAGAGAGTGTGCTTCGANTTACCGAAGCTGAGACTTATAGACATCTCC 1896  
Db 1084 AACACCTTCAGAGAGTGTGCTTCGANTTACCGAAGCTGAGACTTATAGACATCTCC 1143  
Qy 1897 TTCTGCTCCAGC 1909  
Db 1144 TTCTGCTCCAGC 1156

RESULT 11  
US-10-677-943-1  
; Sequence 1, Application US/10677943  
; Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zhi-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT FILING DATE: 2003-10-01  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-677-943-1

Query Match 27.3%; Score 1102; DB 18; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 757 GGTCCGGGAGCGAGATGATGTCGCGACAGAAAGGCTGTGATATGATGACGCTTC 816  
Db 4 GGTCCGGGAGCGAGATGATGTCGCGACAGAAAGGCTGTGATATGATGACGCTTC 63  
Qy 817 GATGACCTGGGCTGTGCTCTCAACAATGACAAAGCTCTGCAAGACTGGGCTGAGAG 876

Db 64 GATGACCTGGGCTGTGCTCTCAACAATGACAAAGCTCTGCAAGACTGGGCTGAGAG 123  
Qy 877 CAGCCTCCGCTTACACCTTCATACAGAGTCTGTAGAGAGGCTCTGCTCCGAGTCTTC 936  
Db 124 CAGCCTCCGCTTACACCTTCATACAGAGTCTGTAGAGAGGCTCTGCTCCGAGTCTTC 183  
Qy 937 CTGATCGTCAACGCTCAGAGAGTGTGGGCAAGAGAGCTCAAGTCAAGAGTGTGTCTCC 996  
Db 184 CTGATCGTCAACGCTCAGAGAGTGTGGGCAAGAGAGCTCAAGTCAAGAGTGTGTCTCC 243  
Qy 997 CGTTACCTGTTAGTTAAGAGATCTCCGGGAAACAAGATTCACCTTGCTCTTGAGCGC 1056  
Db 244 CGTTACCTGTTAGTTAAGAGATCTCCGGGAAACAAGATTCACCTTGCTCTTGAGCGC 303  
Qy 1057 GGGATGCTGAGCATCAAGAAACACAAGAGGTTGGGTGCGATGATGAACAACCGTACCTG 1116  
Db 304 GGGATGCTGAGCATCAAGAAACACAAGAGGTTGGGTGCGATGATGAACAACCGTACCTG 363  
Qy 1117 CTGACACGATGACAGGTGCCCCCGCGTGGGCTCTCATCTGCGTGGCCCTGACCTGAC 1176  
Db 364 CTGACACGATGACAGGTGCCCCCGCGTGGGCTCTCATCTGCGTGGCCCTGACCTGAC 423  
Qy 1177 GACGCTGGGGGAGAGCGTGCCTTCAACCAACCGCTCAGAGGCTTGACGCGCT 1236  
Db 424 GACGCTGGGGGAGAGCGTGCCTTCAACCAACCGCTCAGAGGCTTGACGCGCT 483  
Qy 1237 TTTGTGTTTCAATCAACCCCTCGAGGCGTGGTCCGGGCGTGTCTCATCTGAGAGAA 1296  
Db 484 TTTGTGTTTCAATCAACCCCTCGAGGCGTGGTCCGGGCGTGTCTCATCTGAGAGAA 543  
Qy 1297 AGAGTTGCTGAGAGCGCTTCGCGCTATGAGCTGTGAGAGGAGTGAATAGAGATCA 1356  
Db 544 AGAGTTGCTGAGAGCGCTTCGCGCTATGAGCTGTGAGAGGAGTGAATAGAGATCA 603  
Qy 1357 GTGTTGATGATGACGACCTTCATGTTCAAGAGCTGGGAGTGTGAGCTTCGCTGTG 1416  
Db 604 GTGTTGATGATGACGACCTTCATGTTCAAGAGCTGGGAGTGTGAGCTTCGCTGTG 663  
Qy 1417 TTTCAATGAAACATCTTTCCTCCAGACAGCCACTGTGAGAGATCTACATCTTTCAC 1476  
Db 664 TTTCAATGAAACATCTTTCCTCCAGACAGCCACTGTGAGAGATCTACATCTTTCAC 723  
Qy 1477 CTCAGCTCCAGAGCTCTGAGCGCTTGTACTACGTGTAGAGGCTTGAATTCAG 1536  
Db 724 CTCAGCTCCAGAGCTCTGAGCGCTTGTACTACGTGTAGAGGCTTGAATTCAG 783  
Qy 1537 CCAGCTCTCTGCGCTCTGTACGTGTAGAGACAAAGAGGTCATGAGCTTAAACAGCA 1596  
Db 784 CCAGCTCTCTGCGCTCTGTACGTGTAGAGACAAAGAGGTCATGAGCTTAAACAGCA 843  
Qy 1597 GGCCTTCATATCCACTCGCTTGTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGAC 1656  
Db 844 GGCCTTCATATCCACTCGCTTGTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGAC 903  
Qy 1657 GTAAGAGAGCCCACTGAGAGTCTGTGAGGCTCTCCCGTCCCTGAGGGGTGAAGCAGAG 1716  
Db 904 GTAAGAGAGCCCACTGAGAGTCTGTGAGGCTCTCCCGTCCCTGAGGGGTGAAGCAGAG 963  
Qy 1717 CTTCGACACTGAGGCTCTGTGTTGGGTGAGCAGCCTTAATGCCACCCAGAGACACC 1776  
Db 964 CTTCGACACTGAGGCTCTGTGTTGGGTGAGCAGCCTTAATGCCACCCAGAGACACC 1023  
Qy 1777 CTGACGCGCTTCACACTGTCTTTTTCAGACTCAAGCAAAAGATTGTCGCTTGACATTA 1836  
Db 1024 CTGACGCGCTTCACACTGTCTTTTTCAGACTCAAGCAAAAGATTGTCGCTTGACATTA 1083  
Qy 1837 AACACCTTCAGAGAGTGTGCTTCGANTTACCGAAGCTGAGACTTATAGACATCTCC 1896  
Db 1084 AACACCTTCAGAGAGTGTGCTTCGANTTACCGAAGCTGAGACTTATAGACATCTCC 1143  
Qy 1897 TTCTGCTCCAGC 1909

Db 1144 TTCTGCTCTCAGC 1156

RESULT 12  
US-10-399-443-3  
Sequence 3, Application US/10399443  
Publication No. US20040028669A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as Represented by the  
Secretary, Department of Health & Human Services, The National Institute  
of Health  
APPLICANT: Nelson, Lawrence M.  
APPLICANT: Tong, Zhi-Bin  
APPLICANT: Nelson, Lawrence  
APPLICANT: Zhi-Bin, Tong  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64785  
CURRENT APPLICATION NUMBER: US/10/399,443  
CURRENT FILING DATE: 2003-04-16  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 1075  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-399-443-3

Query Match 16.3%; Score 657; DB 17; Length 1075;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2550 TCTGGCCCTCAGCCCTGTCAGCAACCGAGCTTGACACACCTGCTCTATCCACAACAG 2609  
Db 1 TCTGGCCCTCAGCCCTGTCAGCAACCGAGCTTGACACACCTGCTCTATCCACAACAG 60

Qy 2610 CTTGGGGAACGAGGTGTAAATCTAATCTGTGTGATCTCATGAGCTTCCCACTGTAGTCT 2669  
Db 61 CTTGGGGAACGAGGTGTAAATCTAATCTGTGTGATCTCATGAGCTTCCCACTGTAGTCT 120

Qy 2670 GCAAGAGCTGATGCTGATTCAGTGTGACCTGACACCGCTGCTGTGTTTCTTGCACT 2729  
Db 121 GCAAGAGCTGATGCTGATTCAGTGTGACCTGACACCGCTGCTGTGTTTCTTGCACT 180

Qy 2730 TGCGCTTATGGTAACTCATGAGCTGACGACCTGAGCCTTAGCANTGAACCTGTGGAAGA 2789  
Db 181 TGCGCTTATGGTAACTCATGAGCTGACGACCTGAGCCTTAGCANTGAACCTGTGGAAGA 240

Qy 2790 CAATGGCGTGAAGCTTCTGTGTGAGGTGATGAGAGAACCATCTTGTCTCAAGACCT 2849  
Db 241 CAATGGCGTGAAGCTTCTGTGTGAGGTGATGAGAGAACCATCTTGTCTCAAGACCT 300

Qy 2850 GGAAGTTGTTAAAGTTCATCTCAACCGCGCTGTGTGAGAGTCTGTCTGTGTGATCTC 2909  
Db 301 GGAAGTTGTTAAAGTTCATCTCAACCGCGCTGTGTGAGAGTCTGTCTGTGTGATCTC 360

Qy 2910 GAGGAGGACACCTGGAAGAGCTGATCTGACGGAACAATGCCCGGGGTGACGGTGGGGT 2969  
Db 361 GAGGAGGACACCTGGAAGAGCTGATCTGACGGAACAATGCCCGGGGTGACGGTGGGGT 420

Qy 2970 TGCTGCGCTGTGTGAGGAGTGAAGCAAAAGAACAGTGTCTGAGAGACTCGGGTTGAA 3029  
Db 421 TGCTGCGCTGTGTGAGGAGTGAAGCAAAAGAACAGTGTCTGAGAGACTCGGGTTGAA 480

Qy 3030 GGCATGTGACTGATCTTGAATGCTGTGAGGCACTTCTCTTGGCCCTTCTCTGCAACCG 3089  
Db 481 GGCATGTGACTGATCTTGAATGCTGTGAGGCACTTCTCTTGGCCCTTCTCTGCAACCG 540

Qy 3090 GCATCTGACCAAGTCTGGAAGCTGTGAGAGTAACTTCACTGCCCAAGGAAGATGAAGCT 3149  
Db 3149 GCATCTGACCAAGTCTGGAAGCTGTGAGAGTAACTTCACTGCCCAAGGAAGATGAAGCT 3149

Db 541 GCATCTGACCAAGTCTGGAAGCTGTGAGAGTAACTTCACTGCCCAAGGAAGATGAAGCT 600

Qy 3150 GTGTTGGCCCTTTCGCTGTGCTCCAGCTTAATCAATTAATGGGCTGTGGAATGGCA 3209  
Db 601 GTGTTGGCCCTTTCGCTGTGCTCCAGCTTAATCAATTAATGGGCTGTGGAATGGCA 660

Qy 3210 GTACCTGTGCAATTAAGGAAGCTGTGAGAGTCAAGCTCAAGCCCGAGTCTG 3269  
Db 661 GTACCTGTGCAATTAAGGAAGCTGTGAGAGTCAAGCTCAAGCCCGAGTCTG 720

Qy 3270 AATTGACGGTAAAGTGGCAATCTTTTGAATGAAGATGACCG 3308  
Db 721 AATTGACGGTAAAGTGGCAATCTTTTGAATGAAGATGACCG 759

RESULT 13  
US-10-677-943-3  
Sequence 3, Application US/10677943  
Publication No. US20040072297A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as  
represented by the Secretary of the Department of Health and  
Human Services  
APPLICANT: Nelson, Lawrence  
APPLICANT: Tong, Zhi-Bin  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64790  
CURRENT APPLICATION NUMBER: US/10/677,943  
CURRENT FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US02/09776  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 1075  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-677-943-3

Query Match 16.3%; Score 657; DB 18; Length 1075;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2550 TCTGGCCCTCAGCCCTGTCAGCAACCGAGCTTGACACACCTGCTCTATCCACAACAG 2609  
Db 1 TCTGGCCCTCAGCCCTGTCAGCAACCGAGCTTGACACACCTGCTCTATCCACAACAG 60

Qy 2610 CTTGGGGAACGAGGTGTAAATCTAATCTGTGTGATCTCATGAGCTTCCCACTGTAGTCT 2669  
Db 61 CTTGGGGAACGAGGTGTAAATCTAATCTGTGTGATCTCATGAGCTTCCCACTGTAGTCT 120

Qy 2670 GCAAGAGCTGATGCTGATTCAGTGTGACCTGACACCGCTGCTGTGTTTCTTGCACT 2729  
Db 121 GCAAGAGCTGATGCTGATTCAGTGTGACCTGACACCGCTGCTGTGTTTCTTGCACT 180

Qy 2730 TGCGCTTATGGTAACTCATGAGCTGACGACCTGAGCCTTAGCANTGAACCTGTGGAAGA 2789  
Db 181 TGCGCTTATGGTAACTCATGAGCTGACGACCTGAGCCTTAGCANTGAACCTGTGGAAGA 240

Qy 2790 CAATGGCGTGAAGCTTCTGTGTGAGGTGATGAGAGAACCATCTTGTCTCAAGACCT 2849  
Db 241 CAATGGCGTGAAGCTTCTGTGTGAGGTGATGAGAGAACCATCTTGTCTCAAGACCT 300

Qy 2850 GGAAGTTGTTAAAGTTCATCTCAACCGCGCTGTGTGAGAGTCTGTCTGTGTGATCTC 2909  
Db 301 GGAAGTTGTTAAAGTTCATCTCAACCGCGCTGTGTGAGAGTCTGTCTGTGTGATCTC 360

Qy 2910 GAGGAGGACACCTGGAAGAGCTGATCTGACGGAACAATGCCCGGGGTGACGGTGGGGT 2969  
Db 361 GAGGAGGACACCTGGAAGAGCTGATCTGACGGAACAATGCCCGGGGTGACGGTGGGGT 420

Db 361 GAGGAGCAGACACTGAAGAGCCCTGATCTACGACATATGCTGGTGAAGGTGGGT 420  
QY 2970 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGAGTGTCTGACGAGACTCGGGTTGA 3029  
Db 421 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGAGTGTCTGACGAGACTCGGGTTGA 480  
QY 3030 GGCATGTGAGTGAATCTGTATGTGTGTGAGGACATCTCTGCGCTTCTGCAACCG 3089  
Db 481 GGCATGTGAGTGAATCTGTATGTGTGTGAGGACATCTCTGCGCTTCTGCAACCG 540  
QY 3090 GCATTTGACGAGTAAACCTGTGTGCAAGATTAATCTTCAAGTCCCAAGAGATGTAAGT 3149  
Db 541 GCATTTGACGAGTAAACCTGTGTGCAAGATTAATCTTCAAGTCCCAAGAGATGTAAGT 600  
QY 3150 GTGTGCGGCTTGTGCTGTGCGGAGCTGTAACCTTAAGATTAATGCGGTGGAATGCA 3209  
Db 601 GTGTGCGGCTTGTGCTGTGCGGAGCTGTAACCTTAAGATTAATGCGGTGGAATGCA 660  
QY 3210 GTACCTGTGCAATTAAGAAAGCTGCTGAGGAGTGAAGTCAAGTCAAGCCCGAGTCT 3269  
Db 661 GTACCTGTGCAATTAAGAAAGCTGCTGAGGAGTGAAGTCAAGTCAAGCCCGAGTCT 720  
QY 3270 AATTGACGCTGATGCTGGCATCTTTTGTATGAAGATGACCG 3308  
Db 721 AATTGACGCTGATGCTGGCATCTTTTGTATGAAGATGACCG 759

## RESULT 14

US-10-027-632-258159  
; Sequence 258159, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 258159  
; LENGTH: 2099  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-258159

Query Match 4.3%; Score 174; DB 13; Length 2099;  
Best Local Similarity 100.0%; Pred. No. 7,4e-80;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2000 GGATGCGGATTAAGACCTCTCATTTGAGAGCAGTGGAAAGATTTCTGCTCCATGCTTGGCA 2059  
Db 218 GGATGCGGATTAAGACCTCTCATTTGAGAGCAGTGGAAAGATTTCTGCTCCATGCTTGGCA 277  
QY 2060 CCCACCCACACCTGCGGAGCTGGAAGCTGGGAGAGAGATCCTGACAGAGGGGCGCATGA 2119  
Db 278 CCCACCCACACCTGCGGAGCTGGAAGCTGGGAGAGAGATCCTGACAGAGGGGCGCATGA 337  
QY 2120 AGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAGATACAGACCTGATGT 2173

Db 338 AGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAGATACAGACCTGATGT 391

## RESULT 15

US-10-027-632-258159  
; Sequence 258159, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 258159  
; LENGTH: 2099  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-258159

Query Match 4.3%; Score 174; DB 17; Length 2099;  
Best Local Similarity 100.0%; Pred. No. 7,4e-80;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2000 GGATGCGGATTAAGACCTCTCATTTGAGAGCAGTGGAAAGATTTCTGCTCCATGCTTGGCA 2059  
Db 218 GGATGCGGATTAAGACCTCTCATTTGAGAGCAGTGGAAAGATTTCTGCTCCATGCTTGGCA 277  
QY 2060 CCCACCCACACCTGCGGAGCTGGAAGCTGGGAGAGAGATCCTGACAGAGGGGCGCATGA 2119  
Db 278 CCCACCCACACCTGCGGAGCTGGAAGCTGGGAGAGAGATCCTGACAGAGGGGCGCATGA 337  
QY 2120 AGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAGATACAGACCTGATGT 2173  
Db 338 AGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAGATACAGACCTGATGT 391

Search completed: July 19, 2005, 12:03:16  
Job time : 2599 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2005, 01:37:21 ; Search time 12239 Seconds

(without alignments)  
12549.173 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035  
Sequence: 1 atggaagagagacaatcgct.....tcgaggggcgtgctctctaa 4035

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST.\*

1: gb\_esc1.\*  
2: gb\_esc2.\*  
3: gb\_hnc.\*  
4: gb\_esc3.\*  
5: gb\_esc4.\*  
6: gb\_esc5.\*  
7: gb\_esc6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	679	16.8	791	7	CK000513 AGENCOURT
2	322	8.0	693	5	BU634350 UI-H-FIL-
3	299	7.4	684	5	BU630481 UI-H-FIL-
4	256	6.3	642	5	BU618831 UI-H-FIL-
5	116	2.9	526	8	AO542616 RPTC-11-3
6	60	1.5	523	8	AO428788 CITR-El-
7	58	1.4	814	6	CB228956 AGENCOURT
8	50	1.2	743	9	AG117567 Pan tlogl
9	24	0.6	917	2	BF133827 Reverse s
10	23	0.6	827	7	CR182764 Reverse s
11	22	0.5	308	7	CO111900 EK227159
12	22	0.5	392	7	BE406212 WHE0408.e
13	22	0.5	394	7	CF916020 B0986G02-
14	22	0.5	407	6	CB807101 AMGNNUC:S
15	22	0.5	425	7	BE406757 WHE0406.e
16	22	0.5	511	7	CF173704 B0927G04-
17	22	0.5	533	6	CA559680 K0261H12-
18	22	0.5	539	2	AM948258 RCO-MT001
19	22	0.5	546	6	CA561388 K0286C11-
20	22	0.5	546	7	CK546518 rwbnd_01
21	22	0.5	559	6	CA561556 K0288H05-
22	22	0.5	583	6	CA559979 K0266B01-
23	22	0.5	585	8	AO515975 HS 5240.A
24	22	0.5	589	6	CA559886 K0264H01-

25	22	0.5	593	6	CA558845 K0250F09-
26	22	0.5	599	7	CF915731 B0984E03-
27	22	0.5	605	7	CK516658 rwbnd_00
28	22	0.5	625	6	CD338591 ScFus57.
29	22	0.5	636	7	CF914959 B0971H02-
30	22	0.5	643	7	CF913803 B0954D05-
31	22	0.5	666	1	AV67637 AV367637
32	22	0.5	719	7	CO806156 AGENCOURT
33	22	0.5	768	7	CO808001 AGENCOURT
34	22	0.5	798	2	BB643252 Ct12_8_F0
35	22	0.5	835	4	BI957125 HVSMe000
36	22	0.5	1096	4	BC324013 602422255
37	22	0.5	1119	4	BM454123 AGENCOURT
38	22	0.5	1913	3	AK016782 Mus muscu
39	22	0.5	3475	3	AK087774 Mus muscu
40	22	0.5	395	1	AA316487 EST188400
41	21	0.5	412	6	CB126685 K-EST0175
42	21	0.5	419	7	CN265231 170006000
43	21	0.5	444	2	AM439360 xtl1h09.x
44	21	0.5	456	7	CO380162 FRA3426.S
45	21	0.5	459	1	AI056454 oy77a10.x

#### ALIGNMENTS

RESULT 1  
CK000513  
LOCUS  
DEFINITION AGENCOURT 16368905 NIH MGC 221 Homo sapiens cDNA clone  
IMAGE:30708637 5', mRNA sequence.  
ACCESSION CK000513.1 GI:38526547  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contract: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cga@nci.nih.gov

Tissue Procurement: James Martin, University of Iowa  
cDNA Library Preparation: M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

Plate: NDM1075 row: C column: 14  
High quality sequence stop: 688.  
Location/Qualifiers

#### FEATURES

1..791  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30708637"  
/lab\_host="DH10B Tona"  
/clone\_id="NIH MGC 221"  
/note="Organ: mixed; Vector: pYX-Asc; Site 1: Scori;  
Site 2: NotI; Library is oligo-dT primed and directionally  
cloned. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to cDNA size fraction, ligated with  
EcoR I adaptor, digested with Not I and then cloned  
directionally into pYX-Asc vector. Average insert size  
4-5kb. Adaptors 5' (ATTTCGACACGAG)3' and 5' d

(CCCGCGGCGG) 3', 3' linker sequence - GGGGGCGGCTGAGAGCC T18  
Sequencing primers 3' end: T3 promoter primer 5'd  
(ATTACCCCTCACTAAGAGGA) 3', 5' End: T7 promoter primer 5'd  
(TAATCGACTCACTAAGAGG) 3'. Library was constructed in the  
laboratory of M. Bento Soares. Note: this is a NIH\_MGC  
Library"



```
RESULT 3
BU630481/c
LOCUS      BU630481      684 bp      mRNA      linear      EST 23-SEP-2002
DEFINITION UI-H-FLO-bdk-c-22-0-UI.g1 NCI CGAP FLO Homo sapiens cDNA clone
            UI-H-FLO-bdk-c-22-0-UI 3', mRNA sequence.
ACCESSION  BU630481
VERSION    BU630481.1  GI:23297486
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 684)
            NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
AUTHORS    Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=yes.

FEATURES
            source
            1..684
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-H-FLO-bdk-c-22-0-UI"
            /tissue_type="Cell lines"
            /dev_stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI CGAP FLO"
            /note="Organ: Chondrosarcoma; Vector: pTT73-Pac
            (Pharmacia) with a modified polylinker; Site 1: EcoR I;
            Site 2: Not I; NCI CGAP FLO is a cDNA library derived from
            a pool of mRNA obtained from 4 cell lines from grade III
            chondrosarcoma tissues. The library was constructed
            according to Bonaldo, Lennon and Soares, Genome Research,
            6:791-806, 1996. First strand cDNA synthesis was primed
            with an oligo-dT primer containing a Not I site. Double
            stranded cDNA was ligated to an EcoR I adaptor, digested
            with Not I, and cloned directionally into pTT73-Pac
            vector. The oligonucleotide used to prime the synthesis of
            first-strand cDNA contains a library tag sequence that is
            located between the Not I site and the (dT)18 tail. The
            sequence tag for this library is GAGGTCGGTG. The cell line
            was provided by Dr James Martin from University of Iowa.
            TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
            TAG_LIB=UI-H-FLO
            TAG_SEQ=GAGGTCGGTG"

ORIGIN
Query Match      7.4%; Score 299; DB 5; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.6e-147;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2966 GGGTGTGCGCTGTGCGAGGACGAGCAAGCAAGAGCTTCTGACGAGCTCGGAT 3025
DB 658 GGGTGTGCGCTGTGCGAGGACGAGCAAGCAAGAGCTTCTGACGAGCTCGGAT 599
QY 3026 TGAAGGCACTGACGACTGCTTGTGATGCTGTGAGGCACTCTTGGCCCTTCTCTGA 3085
DB 598 TGAAGGCACTGACGACTGCTTGTGATGCTGTGAGGCACTCTTGGCCCTTCTCTGA 539
QY 3086 ACCGCGATGACGACGCTAAACGCGGTGCAATTAAGTCCCAAGGAATGATGA 3145
|||||
```

```
DB 538 ACCGCGATGACGACGCTAAACCTGTGAGAAATTAAGTCCCAAGGAATGATGA 479
QY 3146 AGCTGTGTTGGCCCTTGGCTTCCCTGCTGCTTAACTTAAGATTAATGGCTGTGGAAT 3205
DB 478 AGCTGTGTTGGCCCTTGGCTTCCCTGCTGCTTAACTTAAGATTAATGGCTGTGGAAT 419
QY 3206 GCGAGTACCTGCTGCAATTAAGAAAGCTGTGAGGAGAGTCACTCAAGCCCGCA 3264
DB 418 GCGAGTACCTGCTGCAATTAAGAAAGCTGTGAGGAGAGTCACTCAAGCCCGCA 360

RESULT 4
BU618831/c
LOCUS      BU618831      642 bp      mRNA      linear      EST 23-SEP-2002
DEFINITION UI-H-FH1-bf1-1-19-0-UI.g1 NCI CGAP FH1 Homo sapiens cDNA clone
            UI-H-FH1-bf1-1-19-0-UI 3', mRNA sequence.
ACCESSION  BU618831
VERSION    BU618831.1  GI:23285046
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 642)
            NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
AUTHORS    Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=yes.

FEATURES
            source
            1..642
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-H-FH1-bf1-1-19-0-UI"
            /tissue_type="Cell line"
            /dev_stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI CGAP FH1"
            /note="Organ: Chondrosarcoma; Vector: pTT73-Pac
            (Pharmacia) with a modified polylinker; Site 1: EcoR I;
            Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library
            obtained from a cell line derived from grade I
            chondrosarcoma tissue. The library was constructed and
            normalized according to Bonaldo, Lennon and Soares, Genome
            Research, 6:791-806, 1996. First strand cDNA synthesis was
            primed with an oligo-dT primer containing a Not I site.
            Double stranded cDNA was ligated to an EcoR I adaptor,
            digested with Not I, and cloned directionally into
            pTT73-Pac vector. The oligonucleotide used to prime the
            synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            (dT)18 tail. The sequence tag for this library is
            AGAATCCGCG. The cell line was provided by Dr. James Martin
            from the University of Iowa.
            TAG TISSUE=Human Chondrosarcoma Cell line C58 - Grade 1
            Chondrosarcoma
            TAG_LIB=UI-H-FH1
            TAG_SEQ=AGAATCCGCG"

ORIGIN
Query Match      6.3%; Score 256; DB 5; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.7e-124;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3009 TCTGACGAGCTCGGGTTGAAGGCGATGTGACTGACTTCTGATTGCTGTGAGGCACTCTC 3068  
 DB 615 TCTGACGAGCACTCGGGTTGAAGGCGATGTGACTGACTTCTGATTGCTGTGAGGCACTCTC 556  
 QY 3069 CTGGCCCTTTCTGCAACCGGCATCTGACCAAGTCTTAACCTGTGCAAGAAATTAATTGAG 3128  
 DB 555 CTGGCCCTTTCTGCAACCGGCATCTGACCAAGTCTTAACCTGTGCAAGAAATTAATTGAG 496  
 QY 3129 TCCCAAGGAATGATGAGCTGTGTGCGCCCTTTGCGCCCTGTCCCAAGTCTTAACAT 3188  
 DB 495 TCCCAAGGAATGATGAGCTGTGTGCGCCCTTTGCGCCCTGTCCCAAGTCTTAACAT 436  
 QY 3189 AATTGGGCTGTGAAATGGCAGTACCTGTGCAATAAGAAAGCTGTGAGAAAGTGA 3248  
 DB 435 AATTGGGCTGTGAAATGGCAGTACCTGTGCAATAAGAAAGCTGTGAGAAAGTGA 376  
 QY 3249 GCTACTCAAGCCCCGA 3264  
 DB 375 GCTACTCAAGCCCCGA 360  
 RESULT 5  
 A0542616/c 526 bp DNA linear GSS 19-MAY-1999  
 LOCUS RPCI-11-345A9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-345A9,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0542616  
 VERSION A0542616.1 GI:4873072  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 526)  
 AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and  
 Venter,J.C.  
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_GSSs: RPCI-11-345A9.TV  
 Contact: Shaying Zhao, William Niernan, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@jlong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/>) or from  
 Research Genet. cs (<http://info@resgen.com>). BAC end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).  
 Seq primer: SP6  
 Class: BAC ends.  
 FEATURES  
 source Location/Qualifiers  
 1..526  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7632104"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-345A9"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /clone\_id="RPCI-11"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"  
 ORIGIN  
 Query Match 2.9%; Score 116; DB 8; Length 526;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-49;  
 Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2516 AGACGTGGCATCAGACCAAGGTTGCCAGAGTCTGGCTCAGCCCTGTCAGCAACC 2575  
 DB 302 AGACGTGGCATCAGACCAAGGTTGCCAGAGTCTGGCTCAGCCCTGTCAGCAACC 243  
 QY 2576 GGACCTTGACACACTTGTGCTTATCCAAACAACGCTGGGGAACGAGGTATAAT 2631  
 DB 242 GGACCTTGACACACTTGTGCTTATCCAAACAACGCTGGGGAACGAGGTATAAT 187  
 RESULT 6  
 A0428788 523 bp DNA linear GSS 24-MAR-1999  
 LOCUS CITBI-EI-2560117.TF CITBI-EI Homo sapiens genomic clone 2560117,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0428788  
 VERSION A0428788.1 GI:4496554  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 523)  
 AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., Shizuya,H., Simon,M. and  
 Venter,J.C.  
 TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready  
 Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_GSSs: CITBI-EI-2560117.TR  
 Contact: Shaying Zhao, William Niernan, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are available from Research Genetics (<http://info@resgen.com>). BAC  
 end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).  
 Seq primer: M13-21  
 Class: BAC ends.  
 FEATURES  
 source Location/Qualifiers  
 1..523  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2560117"  
 /sex="male"  
 /cell\_type="sperm"  
 /clone\_id="CITBI-EI"  
 /note="Vector: pBelOBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
 Caltech Human BAC Library D"  
 ORIGIN  
 Query Match 1.5%; Score 60; DB 8; Length 523;  
 Best Local Similarity 100.0%; Pred. No. 1,3e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 345 AGAATTTTCACACTTATGGAACAAGAGTCCACGACGACAGACAGAAACAAG 404  
 DB 21 AGAATTTTCACACTTATGGAACAAGAGTCCACGACGACAGACAGAAACAAG 80  
 RESULT 7  
 CB228956 814 bp mRNA linear EST 10-FEB-2003  
 LOCUS AGENCOURT.11501187 NICHD\_Rh\_OVI Macaca mulatta cDNA clone  
 DEFINITION IMAGE:6884760 5', mRNA sequence.  
 ACCESSION CB228956  
 VERSION CB228956.1 GI:28280534  
 KEYWORDS EST.  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM Macaca mulatta

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca. 1 (bases 1 to 814)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Elliot Spindel  
cDNA Library Preparation: CLONTECH  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1CM3135 row: b column: 23  
High quality sequence stop: 235.  
Location/Qualifiers  
1. 814  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/db\_xref="taxon:9544"  
/clone="IMAGE:6884760"  
/tissue\_type="Ovary"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHD\_Rh\_OVI"  
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I; Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.0-4.0 kb. Tissue pooled from pre-pubertal, post pubertal sn menopausal monkeys. Constructed by Clontech. Note: this is a NICHD Library."  
ORIGIN  
Query Match 1.4%; Score 58; DB 6; Length 814;  
Best Local Similarity 100.0%; Pred. No. 1.6e-18;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2289 GATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGCTTTGAGCGCTGAT 2346  
140 GATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGCTTTGAGCGCTGAT 197  
RESULT 8  
LOCUS AG117567 743 bp DNA linear GSS 03-NOV-2001  
DEFINITION Pan troglodytes DNA, clone: PTB-125H04.F, genomic survey sequence.  
ACCESSION AG117567  
VERSION AG117567.1 GI:16738086  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.  
REFERENCE 1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
BAC end sequences of library PTB  
Unpublished  
2 (bases 1 to 743)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: [chumpeesgsc.riken.go.jp](mailto:chumpeesgsc.riken.go.jp), URL: <http://hgp.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  
PRIMERS  
COMMENT

Sequencing: -21M13  
LIBRARY Vector : PKS145  
R.Site 1 : SacI  
R.Site 2 : SacI.  
Location/Qualifiers  
1. 743  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-125H04.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
ORIGIN  
Query Match 1.2%; Score 50; DB 9; Length 743;  
Best Local Similarity 100.0%; Pred. No. 3e-14;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3332 GCGTCTCTGAAAGCCGGCATGCGCATGTGCTTCTGTGGGATGAC 3381  
296 GCGTCTCTGAAAGCCGGCATGCGCATGTGCTTCTGTGGGATGAC 247  
RESULT 9  
LOCUS BF133827 917 bp mRNA linear EST 24-OCT-2000  
DEFINITION 601778752P1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4006716.5', mRNA sequence.  
ACCESSION BF133827  
VERSION BF133827.1 GI:10972867  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 917)  
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>. National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM9238 row: h column: 13  
High quality sequence stop: 630.  
Location/Qualifiers  
1. 917  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4006716"  
/tissue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu30"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model MMT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"  
ORIGIN  
Query Match 0.6%; Score 24; DB 2; Length 917;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1895 CCTTCTGCTCCAGCACTGTCCGT 1918
      |||||||
Db      742 CCTTCTGCTCCAGCACTGTCCGT 719

RESULT 10
LOCUS   CR182764                827 bp    DNA        linear    GSS 06-JUL-2004
DEFINITION Reverse strand read from insert in 5'HPT insertion targeting and
            chromosome engineering clone MHPN374002, genomic survey sequence.
ACCESSION CR182764
VERSION   CR182764.1 GI:49961613
KEYWORDS  GSS; genome survey sequence; MICR.
SOURCE    Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 827)
            Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,D.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
TITLE      Direct Submission
JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICR
FEATURES
            source
            1..827
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHPN374002"
            /clone_lib="MHPN"

ORIGIN
Query Match      0.6%; Score 23; DB 9; Length 827;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2977 CTGTGCGAGGAGCTAGAGCAAAA 2999
      |||||||
Db      49 CTGTGCGAGGAGCTAGAGCAAAA 71

RESULT 11
LOCUS   CO311900/c              308 bp    mRNA        linear    EST 28-JUN-2004
DEFINITION EK227159.5prime Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
            melanogaster cDNA clone EK227159 5, mRNA sequence.
ACCESSION CO311900
VERSION   CO311900.1 GI:49366701
KEYWORDS  EST.
SOURCE    Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 308)
            Kopczynski,C., Platt,D., Campbell,J., Muzong,C., Laufer,A.,
            Peterson,E. and Swimmer,C.
            Exelixis FlyTag EST Project CK01 Library
            Unpublished (2004)
            Contact: Stapleton, M.
TITLE      BDGP
JOURNAL   Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: EK-2271 row: B column: 11
            High quality sequence stop: 267.
FEATURES
            source
            1..308
            /organism="Drosophila melanogaster"
            /mol_type="mRNA"
            /db_xref="taxon:7227"

```

```

ORIGIN
Query Match      0.5%; Score 22; DB 7; Length 308;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2553 GGCTCAGCCCTGCTCAGCAAC 2574
      |||||||
Db      201 GGCTCAGCCCTGCTCAGCAAC 180

RESULT 12
LOCUS   BE406212                392 bp    mRNA        linear    EST 21-JUL-2000
DEFINITION WHE0408_e11_e12B wheat etiolated seedling root cDNA library
            Triticum aestivum cDNA clone WHE0408_e11_e11, mRNA sequence.
ACCESSION BE406212
VERSION   BE406212.1 GI:9365680
KEYWORDS  EST.
SOURCE    Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooidae; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 392)
            Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D.,
            Han,P.S., Heia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rauech,C.J.,
            Seaton,C.L. and Tong,J.C.
            The structure and function of the expressed portion of the wheat
            genomes
            Unpublished (2000)
            Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Strategene pBluescript SK reverse primer.
FEATURES
            source
            1..392
            /organism="Triticum aestivum"
            /mol_type="mRNA"
            /cultivar="Chinese Spring"
            /db_xref="taxon:4565"
            /clone="WHE0408_e11_e11"
            /tissue_type="Root"
            /dev_stage="Five day old etiolated seedling"
            /lab_host="E. coli SOLR"
            /clone_lib="wheat etiolated seedling root cDNA library"
            /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
            Site_1: EcoRI; Site_2: XhoI; Seeds were
            surface-sterilized, germinated and grown aseptically in
            the dark at room temperature on filter paper with water,
            nystatin and cefotaxime in covered crystallization
            dishes. Roots were harvested. The tissue, total RNA, and
            poly(A) RNA were prepared, a cDNA library was made, and
            the cDNA clones were in vivo excised to give plusscript
            phagemids in the T7 Close lab (Choi, Close, Fenton) at the
            University of California, Riverside. Plasmid DNA
            preparations and DNA sequencing were performed in the OD
            Anderson lab (all other authors)."
ORIGIN
Query Match      0.5%; Score 22; DB 2; Length 392;

```

Best Local Similarity 100.0%; Pred. No. 28;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1482 TTCCAGAGACTTCTGTGCGGCC 1503  
|||||  
162 TCTCCAGAGACTTCTGTGCGGCC 183

## RESULT 13

CF916020

LOCUS CF916020 394 bp mRNA linear EST 05-NOV-2003

DEFINITION B0988G02-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus

musculus cDNA clone NIA:B0988G02 IMAGE:30480745 5', mRNA sequence.

ACCESSION CF916020

VERSION CF916020.1 GI:38187222

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 394) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Piao, Y., Ko, N.-I., Lim, M. K. and Ko, M. S. H. Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL 21429098

MEDLINE 11544199

COMMENT Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: B0988 row: G column: 02

Seq primer: M13 Reverse

High quality sequence stop: 394

POLYA=No.

FEATURES

source

1. 394

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/db\_xref="taxon:10090"

/clone="NIA:B0988G02 IMAGE:30480745"

/dev\_stage="Unfertilized Egg"

/lab\_host="DH10B"

/clone\_1lb="NIA Mouse Unfertilized Egg cDNA Library (Long

1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site\_1: SalI;

Site\_2: NotI; Mouse cDNA project by the Laboratory of

Genetics, National Institute on Aging (NIA), Intramural

Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).

This is a long-transcript enriched cDNA library (Ref.

Genome Res. 11: 1553-1558 (2001). (PMID: 11544199). Total

RNA were extracted from a pool of 1468 unfertilized eggs.

Double-stranded cDNAs were synthesized with an Oligo(dT)

primer (Invitrogen):

5'-PACATGATCTAGATCGAGCGCGCCCTTTT-3'.

treated with T4 DNA polymerase, and purified by

ethanol-precipitation. The cDNAs were ligated to

lone-linker L1-SalI, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Tag polymerase (Takara) with a primer SalI-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes

and cloned into SalI/NotI site of pCMV-SPORT6 plasmid

vector. The DH10B E. coli host was transformed with the

ligation mixture by the standard chemical method. The

average insert size is about 2.5 kb. The library was

constructed by Yulan Piao."

Query Match 0.5%; Score 22; DB 7; Length 394;

Best Local Similarity 100.0%; Pred. No. 28;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1417 TTTCATGATGATCTCTCTCC 1438  
|||||  
56 TTTCATGATGATCTCTCTCC 77

## RESULT 14

CB807101

LOCUS CB807101 407 bp mRNA linear EST 16-MAY-2003

DEFINITION AMGNNUC:SRPB2-00129-B10-A srpb2 (10220) Rattus norvegicus cDNA

clone srpb2-00129-b10 5', mRNA sequence.

ACCESSION CB807101

VERSION CB807101.1 GI:29925806

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 407) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

Amgen EST Program.

Amgen Rat EST Program

Unpublished (2003)

CONTACT: Dan Fitzpatrick

Amgen, Inc

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00129 row: b column: 10.

FEATURES

source

1. 407

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="srpb2-00129-b10"

/tissue\_type="prostate tissue"

/clone\_1lb="srpb2 (10220)"

/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; rat

prostate normalized double selected poly(A+) mRNA size

fraction > 1 kb"

ORIGIN

Query Match 0.5%; Score 22; DB 6; Length 407;

Best Local Similarity 100.0%; Pred. No. 28;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 482 TTGAACACATGCTGCTGACTG 503  
|||||  
167 TTGAACACATGCTGCTGACTG 188

RESULT 15

BE406757

LOCUS BE406757 425 bp mRNA linear EST 21-JUL-2000

DEFINITION WHB0406 e11 e11z wheat etiolated seedling root cDNA library

Triticum aestivum cDNA clone WHB0406\_e11\_e11, mRNA sequence.

ACCESSION BE406757

VERSION BE406757.1 GI:9366225

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 425) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Pooideae; Triticeae; Triticum.

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,

Han, P.S., Hsieh, C.C., Kang, Y., Lazo, G.R., Miller, R., Rauech, C.J.,

Seaton, C.L. and Tong, J.C.

The structure and function of the expressed portion of the wheat

genomes

ORIGIN

JOURNAL  
COMMENT

Unpublished (2000)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanderson@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: Strategene SK primer.

## FEATURES

source

1. .425  
 location/Qualifiers  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE0406\_e11\_e11"  
 /tissue\_type="Root"  
 /dev\_stage="Five day old etiolated seedling"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="wheat etiolated seedling root cDNA library"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Seeds were  
 surface-sterilized, germinated and grown aseptically in  
 the dark at room temperature on filter paper with water,  
 nystatin and cefotaxime in covered crystallization  
 dishes. Roots were harvested. The tissue, total RNA, and  
 poly(A) RNA were prepared, a cDNA library was made, and  
 the cDNA clones were in vivo excised to give phagescript  
 phagemids in the TJ Clonase lab (Choi, Close, Fenton) at the  
 University of California, Riverside. Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

## ORIGIN

Query Match 0.5%; Score 22; DB 2; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1482 TCTCCAGGACTTCTGTGCGGCC 1503  
 ||||||||||||||||||||  
 Db 117 TCTCCAGGACTTCTGTGCGGCC 138

Search completed: July 19, 2005, 11:09:20  
 Job time : 12248 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 14:43:30 ; Search time 17004 Seconds  
(without alignments)  
11498.283 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atggaagagagaacacgcct.....tcgaggcgctggtcctaa 4035

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_srs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3172.6	78.6	3885	9	AY154460 Homo sapi
2	3166.2	78.5	3885	9	AY054986 Homo sapi
3	3166.2	78.5	3900	6	AX427610 Sequence
4	3139	77.8	3830	6	AX704823 Sequence
5	3133.6	77.7	3489	6	AX478549 Sequence
6	3133.6	77.7	3926	6	AX704821 Sequence
7	3057.8	75.8	5859	6	AX459881 Sequence
8	3057.8	75.8	6939	6	AX459873 Sequence
9	3057.8	75.8	6939	6	AX459891 Sequence
10	2553.8	63.3	2753	6	CQ731113 Sequence
11	1598.8	39.6	157141	9	AC011470 Homo sapi
12	1598.8	39.6	193609	9	AC024580 Homo sapi
13	1566	38.8	3545	4	AY721594 Bos tauru
14	1287.8	31.9	3405	10	AY329487 Mus muscu
15	1286.2	31.8	3432	10	AY329484 Mus muscu
16	1282.8	31.8	3480	10	AY196361 Mus muscu
17	1281.2	31.7	3470	10	AY196362 Mus muscu
18	1279.6	31.7	3453	10	AY329486 Mus muscu
19	1278	31.7	3447	6	AX427592 Sequence

20	1278	31.7	3447	10	AF074018 Mus muscu
21	1278	31.7	3454	10	AY329490 Mus muscu
22	1276.4	31.6	3534	10	BC053384 Mus muscu
23	1152.8	28.6	1157	6	AX427588 Sequence
24	1144	28.4	3375	10	AY329489 Mus muscu
25	1140.8	28.3	3372	10	AY329485 Mus muscu
26	1008.6	25.0	3303	10	AY329488 Mus muscu
27	1005.4	24.9	3303	10	AY329491 Mus muscu
28	988.6	24.5	167509	2	AC012107 Homo sapi
29	840.4	20.8	193662	2	AC127352 Mus muscu
30	840.4	20.8	213424	2	AC148980 Mus muscu
31	840.4	20.8	299377	2	AC127238 Mus muscu
32	756.6	18.8	1075	6	AX427590 Sequence
33	729.4	18.1	3282	9	AY154469 Homo sapi
34	649.4	16.1	231371	2	AC132633 Rattus no
35	649.4	16.1	323826	2	AC110429 Rattus no
36	645.2	16.0	1576	10	AF143565 Mus muscu
37	528.4	13.1	167509	2	AC012107 Homo sapi
38	430.4	10.7	2964	10	AY596199 Mus muscu
39	428.8	10.6	3281	10	AY673647 Mus muscu
40	422.4	10.5	3190	6	AX714395 Sequence
41	422.4	10.5	3190	9	AK056688 Homo sapi
42	422.4	10.5	3333	9	AF442488 Homo sapi
43	422.4	10.5	3339	9	AY072792 Homo sapi
44	422.4	10.5	3368	9	AF479747 Homo sapi
45	419.2	10.4	2575	6	AX320139 Sequence

## ALIGNMENTS

RESULT 1	AY154460	3885 bp	mRNA	linear	PRI 20-FEB-2003
LOCUS	AY154460				
DEFINITION	Homo sapiens NALP5 (NALP5) mRNA, complete cds.				
ACCESSION	AY154460				
VERSION	AY154460.1 GI:28436363				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3885)				
AUTHORS	Tschopp,J., Martinon,F. and Burns,K.				
JOURNAL	NALP5: a novel protein family involved in inflammation				
MEDLINE	Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)				
PUBMED	12563287				
REFERENCE	2 (bases 1 to 3885)				
AUTHORS	Martinon,F., Hotmann,K. and Tschopp,J.				
TITLE	Direct Submissio				
JOURNAL	Submitted (25-SEP-2002) Institute of Biochemistry, University of				
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland					
FEATURES					
source	1..3885				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
gene	1..3885				
	/gene="NALP5"				
CDS	1..3603				
	/gene="NALP5"				
	/note="member of the NALP protein family involved in				
	inflammation; contains NACHT, LRR and PYD domains; similar				
	to mouse MATER protein"				
	/codon_start=1				
	/product="NALP5"				
	/protein_id="AA018156.1"				
	/db_xref="GI:28436364"				
	/translation="MKVAGGELGAAALSLSPALVTLSTGPTCSILPKNPDPONT				
	SSQCIKMGKSLTFSYGLWQICVLEDEKKEPOTFEKLKKSSSTGTCISIPQFEIE				
	NANVECALILHRYVAGSLAATATSTISFENNNTLTLSRKADDMKRRSPDPKATMD				
	QGPSKEKVPGISQAVQODSATRAETKEDEISQANQEGATATLETBQETISQANQEGA				

TAATEBOGHGDTWYKSHVMTKEAREEDVARSFENTADMPMEMQTLAAGFDSNRWG  
FRPRTVLHSGKSGKLSALARIYLCMAQGLYOCMFYFLPREKORKESSVTE  
FISREMPDQSAVPEIMRSPRILFIIDGPDIGSVLNDPRKCDMAEKOPPTLIR  
SLARKVLIPSEFLIYTVRDVGTREKKSVEPRYLIVGISEGRIHLLRGGEGHO  
KIQGLATMNNBELDQCVPAVSLICVALQLODVESVAPFQIITGLHAAVPH  
QUTPRGVARCINLEERVVLKRFCEMAVEGWNKRSVADGLWQVGGSEBLAFH  
NMLILPSHCEBYTFPHSLIDPCALUYVEGLEIPALCPVEKTKSMELKOA  
GFHLSHMKERFLFGLVSEDEVARPLEVILGCEVPLGVOKLHWVSLIGQOPNLTTR  
DPLDAFHLFETODEFVRLALNSFOEVLPIPNOMLIIASSFCIOHPYLRKIRVDY  
KGIIPRDESAECAPVPLMWRDKLIREOMDPCSMGTHHLOLDIGSSILIRBAM  
KTLCAKLEHPITCKIOTLMPRNAQITPVGQHMRIYMANRNLRLMLGTHLKEEDVRH  
ACEALRHPKCLLESRLDCGGLTHACVLIKISQIILTSLSLSIAGKRVDTGEMPL  
SDALRVSCQALSLIEDCGIATGCGSLAGALVSNRSLTHLSJNPNEDNGVLLCE  
RSMRLPHCSLORLMOCHLDTPAGCGFALAMGSMYTLTHLSJNPNEDNGVLLCE  
VMRPSCHLODELVKCHLTAAACCELSCVLSRSHKLSJLTDNALGDGVALLCEG  
LKOKNSVLTGLKACGLTSDCCERLSLALSCLRHTLSJLNVONNPSKGMKCSAF  
ACPTSNLOIIGLMKWQIPVOIRKLEBVLKPRVLDGSMHSFDEDRYMKW"

## ORIGIN

Query Match 78.6%; Score 3172.6; DB 9; Length 3885;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 123; Gaps 1;

QY 1 ATGGAAGAGACAAATCGCTCACTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60  
DB 154 ATGGAAGAGACAAATCGCTCACTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 213  
QY 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAATAAGAAATTTTCAGAAATCG 120  
DB 214 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAATAAGAAATTTTCAGAAATCG 273  
QY 121 ACCACATGCTCTATTCACAGATTTGAATTCAGAAATGCAACGTGGAATGTCTGGCACTC 180  
DB 274 ACCACATGCTCTATTCACAGATTTGAATTCAGAAATGCAACGTGGAATGTCTGGCACTC 333  
QY 181 CTCTTGATGAGATTTATGAGAGCATGCTGGCTTGGGCTACGTCATTAGCATTTTGA 240  
DB 334 CTCTTGATGAGATTTATGAGAGCATGCTGGCTTGGGCTACGTCATTAGCATTTTGA 393  
QY 241 AACATGAACCTGCCAACCTCTCGGAGAAAGCAAGGATGACATGAAAA----- 289  
DB 394 AACATGAACCTGCCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAATTCACCA 453  
QY 290 ----- 289  
DB 454 GAAGATCCTGAAGCAAGATGACTGACCAAGGACCAAGCAAGAAAAATGCCAGGAATT 513  
QY 290 -----AATTTTCA 297  
DB 514 TCACAAAGCTGTGCAACAAAGATAGTGCAACGCTGCAAGACAAAGAAATTTTCA 573  
QY 298 CAAGCTATGGAACAAAGAGTGCCACAGACGACAGACAAAGAAATTTTCA 357  
DB 574 CAAGCTATGGAACAAAGAGTGCCACAGACGACAGACAAAGAAATTTTCA 633  
QY 358 GCTATGGAACAAGAGTGCCACAGACGACAGACAAAGAAATTTTCA 417  
DB 634 GCTATGGAACAAGAGTGCCACAGACGACAGACAAAGAAATTTTCA 693  
QY 418 ACATGGGACCTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGTAAGTCTG 477  
DB 694 ACATGGGACCTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGTAAGTCTG 753  
QY 478 AGTTTGAACCACTGCTGCTGACCTGGCCGGAATATGCAACGTTGGCTGTGCTTTGAT 537  
DB 754 AGTTTGAACCACTGCTGCTGACCTGGCCGGAATATGCAACGTTGGCTGTGCTTTGAT 813  
QY 538 TCAGACCGGTGGGGCTTCCGGCTTCGACAGGTGTTCTGACAGAAATTCAGAAATTTGG 597  
DB 814 TCAGACCGGTGGGGCTTCCGGCTTCGACAGGTGTTCTGACAGAAATTCAGAAATTTGG 873  
QY 598 AAATCGCTTACGCAAGAGATCGTGTGCTGGGGCGCAAGGTGACTTACACAGGGA 657

DB 874 AAATCGCTTACGCAAGAGATCGTGTGCTGGGGCGCAAGGTGACTTACAGGGA 933  
QY 658 ATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGATGACGGGAAAGAGAGAGAGT 717  
DB 934 ATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGATGACGGGAAAGAGAGAGAGT 993  
QY 718 GTCAAGAGTTCACTCCAGAGAGTGGCCAGACTCCAGAGCTCCGATGACGAGATCATG 777  
DB 994 GTCAAGAGTTCACTCCAGAGAGTGGCCAGACTCCAGAGCTCCGATGACGAGATCATG 1053  
QY 778 TCCGACCAAGAAAGCTGTTTTCATCATTTGACGCTTTCCATGACCTGGGCTCTGCTC 837  
DB 1054 TCCGACCAAGAAAGCTGTTTTCATCATTTGACGCTTTCCATGACCTGGGCTCTGCTC 1113  
QY 838 AACATGACAAAGAGCTGTGAAAGACTGGGCTGAGAGAGCCTCCGTTACACCTGATA 897  
DB 1114 AACATGACAAAGAGCTGTGAAAGACTGGGCTGAGAGAGCCTCCGTTACACCTGATA 1173  
QY 898 CGCAGTCTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACCGTACAGAGAC 957  
DB 1174 CGCAGTCTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACCGTACAGAGAC 1233  
QY 958 GTGGGACAGAGAGCTCAAGTCAAGAGTGTGCTTCCCGTTACCTGTTAGTAGAGA 1017  
DB 1234 GTGGGACAGAGAGCTCAAGTCAAGAGTGTGCTTCCCGTTACCTGTTAGTAGAGA 1293  
QY 1018 ATCTCCCGGGAACAAAGATTCACATTCCTCTGAGGCGGGAATGGTAGACATCAAG 1077  
DB 1294 ATCTCCCGGGAACAAAGATTCACATTCCTCTGAGGCGGGAATGGTAGACATCAAG 1353  
QY 1078 ACAAGAGGTTGCTGAGATGATGAACAACGTGAGCTGCTGACAGAGTCCAGAGTCCC 1137  
DB 1354 ACAAGAGGTTGCTGAGATGATGAACAACGTGAGCTGCTGACAGAGTCCAGAGTCCC 1413  
QY 1138 GCCGTGGCTCTCTCATCTGCTGGGCTCTGACGCTGACGAGCTGTGGGGAGAGGCTC 1197  
DB 1414 GCCGTGGCTCTCTCATCTGCTGGGCTCTGACGCTGACGAGCTGTGGGGAGAGGCTC 1473  
QY 1198 GCCCCTTCAACAAAGGCTCAGAGGCTGACGCGCTTTTGTGTTCAATCACTGACC 1257  
DB 1474 GCCCCTTCAACAAAGGCTCAGAGGCTGACGCGCTTTTGTGTTCAATCACTGACC 1533  
QY 1258 CCTCGAGGCTGTGCTCGGCGCTGTCTCAATCTGAGAGAAAGATGTCTGAAAGCGCTTC 1317  
DB 1534 CCTCGAGGCTGTGCTCGGCGCTGTCTCAATCTGAGAGAAAGATGTCTGAAAGCGCTTC 1593  
QY 1318 TGCCGTATGCTGTGAGAGAGTGTGAATGAGAAATCAAGTGTGATGAGAGACCTC 1377  
DB 1594 TGCCGTATGCTGTGAGAGAGTGTGAATGAGAAATCAAGTGTGATGAGAGACCTC 1653  
QY 1378 ATGCTTCAAGGATCTGGGAGATCTGAGCTCCGTCTCTGTTTCAATGAAACATCTTCTC 1437  
DB 1654 ATGCTTCAAGGATCTGGGAGATCTGAGCTCCGTCTCTGTTTCAATGAAACATCTTCTC 1713  
QY 1438 CCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCACTGACAGATCTTCTGT 1497  
DB 1714 CCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCACTGACAGATCTTCTGT 1773  
QY 1498 GCCCCTTGTATCAAGTGTGAGAGGCTGGAATTCAGGCAAGCTCTGTCCTCTGAC 1557  
DB 1774 GCCCCTTGTATCAAGTGTGAGAGGCTGGAATTCAGGCAAGCTCTGTCCTCTGAC 1833  
QY 1558 GTTGAAGAGCAAGAGTCTCATGAGACTTAAACAGGCAAGCTTCCATATCACTGCTT 1617  
DB 1834 GTTGAAGAGCAAGAGTCTCATGAGACTTAAACAGGCAAGCTTCCATATCACTGCTT 1893  
QY 1618 TGAATGAAGGTTCTTGTGCTGTGAGAGCAAGTGAAGAGGCACTGAGAGGTC 1677  
DB 1894 TGAATGAAGGTTCTTGTGCTGTGAGAGCAAGTGAAGAGGCACTGAGAGGTC 1953  
QY 1678 CTGCTGGGCTGTCCCTTCTTCCCTGTGGGAGTGAAGAGAGCTTCTGCACTGAGTCTCTG 1737  
DB 1954 CTGCTGGGCTGTCCCTTCTTCCCTGTGGGAGTGAAGAGAGCTTCTGCACTGAGTCTCTG 2013

1738 TTGGGTCAAGACCTAATGCAACCAACCCAGAGAGACACCTGAGCCCTTCCACTGTCTT 1797  
 2014 TTGGGTCAAGACCTAATGCAACCAACCCAGAGAGACACCTGAGCCCTTCCACTGTCTT 2073  
 1798 TTGAGACTCAAGACCAAGAGATTGTGTGCTTGGCAATTAACAGCTTCCAGAAAGTGTGG 1857  
 2074 TTGAGACTCAAGACCAAGAGATTGTGTGCTTGGCAATTAACAGCTTCCAGAAAGTGTGG 2133  
 1858 CTTCCGATTAAACCAAGACCTGAGACTTGTAGATCTTCTTGTGCTTCCAGAGACTGTCTGG 1917  
 2134 CTTCCGATTAAACCAAGACCTGAGACTTGTAGATCTTCTTGTGCTTCCAGAGACTGTCTGG 2193  
 1918 TATTGGCGGAAAAATTCGGGTGAGATGTCAAAGGAGATCTTCCAAAGAGATGAGCCGCTGAG 1977  
 2194 TATTGGCGGAAAAATTCGGGTGAGATGTCAAAGGAGATCTTCCAAAGAGATGAGCCGCTGAG 2253  
 1978 GCATGTCCTGTGTGCTCTCTATGAGATGCGGGATTAAGACCTTCATTTAGAGAGAGAGTGGAA 2037  
 2254 GCATGTCCTGTGTGCTCTCTATGAGATGCGGGATTAAGACCTTCATTTAGAGAGAGAGTGGAA 2313  
 2038 GATTTCGTGCTCATGCTTGGACACCAACCAACCTGCGGAGCTGAGACTTGGGAGAGAC 2097  
 2314 GATTTCGTGCTCATGCTTGGACACCAACCAACCTGCGGAGCTGAGACTTGGGAGAGAC 2373  
 2098 ATCCGAGAGAGCGGGCCATGAGAACCTGTGTGCGCAAGCTGAGGAGATCCCACTGAGAG 2157  
 2374 ATCCGAGAGAGCGGGCCATGAGAACCTGTGTGCGCAAGCTGAGGAGATCCCACTGAGAG 2433  
 2158 ATACAGACCTGATTTTGAAGATGACAGATTAACCTGCTGTGTGAGAGACTTGTGAGAG 2217  
 2434 ATACAGACCTGATTTTGAAGATGACAGATTAACCTGCTGTGTGAGAGACTTGTGAGAG 2493  
 2218 ATGCTATGCGCAACCGTAAACCTTAAGATCCCTCAACTTGGAGAGACCACTGAGAGAA 2277  
 2494 ATGCTATGCGCAACCGTAAACCTTAAGATCCCTCAACTTGGAGAGACCACTGAGAGAA 2553  
 2278 GAGAGATGTAAGATGCGGTGAGAACCTTAAACCAACCAATGTTTGTGAGAGCTTTTG 2337  
 2554 GAGAGATGTAAGATGCGGTGAGAACCTTAAACCAACCAATGTTTGTGAGAGCTTTTG 2613  
 2338 AGGCTGATGTTGCTGTGATTTGACCAATGCTGTGCTGAGAGATCTCCCAATCTTACG 2397  
 2614 AGGCTGATGTTGCTGTGATTTGACCAATGCTGTGCTGAGAGATCTCCCAATCTTACG 2673  
 2398 ACCCTCCCAAGCTGAAATCTTCTGAGCTTGGCAGAGAAACAAGTGAACAGACGAGAGTA 2457  
 2674 ACCCTCCCAAGCTGAAATCTTCTGAGCTTGGCAGAGAAACAAGTGAACAGACGAGAGTA 2733  
 2458 ATGCTCTCAAGTATGCTTGAAGATCTTCCAGTGTGCGCTTGCAGAGACTGATCTGAG 2517  
 2734 ATGCTCTCAAGTATGCTTGAAGATCTTCCAGTGTGCGCTTGCAGAGACTGATCTGAG 2793  
 2518 GACTGTGAGATCAAGACCGGTTGCGAGATGCTGAGCTTGAAGCTGCTGCAAGACCGG 2577  
 2794 GACTGTGAGATCAAGACCGGTTGCGAGATGCTGAGCTTGAAGCTGCTGCAAGACCGG 2853  
 2578 AGCTTGACACACTGTGCTATTCACAAACAAGCTGAGGAGAGAGTGAATCTACTG 2637  
 2854 AGCTTGACACACTGTGCTATTCACAAACAAGCTGAGGAGAGAGTGAATCTACTG 2913  
 2638 TGTGATCAATGAGGCTTCCCACTGTAGTCTGCAAGAGCTGATCTGAATCAAGTCCAC 2697  
 2914 TGTGATCAATGAGGCTTCCCACTGTAGTCTGCAAGAGCTGATCTGAATCAAGTCCAC 2973  
 2698 CTGACAGAGGCTGCTGTGTTTCTTGAACCTTGTGAGTGAAGTGAATCAAGTCCAC 2757  
 2974 CTGACAGAGGCTGCTGTGTTTCTTGAACCTTGTGAGTGAAGTGAATCAAGTCCAC 3033  
 2758 CACCTGAGCTTTAGATGAACCTGTGAGAGAAACAATGAGCTGATCTGTGCGAGGTC 2817  
 3034 CACCTGAGCTTTAGATGAACCTGTGAGAGAAACAATGAGCTGATCTGTGCGAGGTC 3093

2818 ATGAGAGAACCATCTTGTATCTCCAGAGACCTGAGATTGTGAAGTGTATCTACACGCC 2877  
 3094 ATGAGAGAACCATCTTGTATCTCCAGAGACCTGAGATTGTGAAGTGTATCTACACGCC 3153  
 2878 GGTGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGACAGACCTGAAGAGCTTGAT 2937  
 3154 GGTGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGACAGACCTGAAGAGCTTGAT 3213  
 2938 CTCACGAGCAATGCTGAGGAGAGAGAGTGTGAGTGTGAGAGAGAGAGAGAGAGAG 2997  
 3214 CTCACGAGCAATGCTGAGGAGAGAGAGTGTGAGTGTGAGAGAGAGAGAGAGAGAG 3273  
 2998 AAGAACAGTGTGTGAG 3057  
 3274 AAGAACAGTGTGTGAG 3333  
 3058 GAGGACCTCTCTTGGCCCTTCTGCAACCGGACCTGACCAAGTCTAACTGAGTGGAG 3117  
 3334 GAGGACCTCTCTTGGCCCTTCTGCAACCGGACCTGACCAAGTCTAACTGAGTGGAG 3393  
 3118 AATACTTCACTCCCAAG 3177  
 3394 AATACTTCACTCCCAAG 3453  
 3178 AACTTCACTTCACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3237  
 3454 AACTTCACTTCACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3513  
 3238 GAGAGAGTGAAGTCACTCAACCGGAGTGTGATTTGAGAGAGAGAGAGAGAGAGAG 3297  
 3514 GAGAGAGTGAAGTCACTCAACCGGAGTGTGATTTGAGAGAGAGAGAGAGAGAGAG 3573  
 3298 GAGAGTGAAGTCACTCAACCGGAGTGTGATTTGAGAGAGAGAGAGAGAGAGAG 3357  
 3574 GAGAGTGAAGTCACTCAACCGGAGTGTGATTTGAGAGAGAGAGAGAGAGAGAG 3633

RESULT 2  
 AT054986  
 LOCUS  
 DEFINITION  
 Homo sapiens maternal-antigen-thal-embryos-require protein (MATER)  
 ACCESSION  
 AY054986  
 VERSION  
 AY054986.1 GI:19882272  
 KEYWORDS  
 ORGANISM  
 Homo sapiens (human)  
 SOURCE  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 3885)  
 Tong, Z.-B., Bondy, C.A., Zhou, J., and Nelson, L.M.  
 A human homologue of mouse Mater, a maternal effect gene essential  
 for early embryonic development  
 Hum. Reprod. 17 (4), 903-911 (2002)  
 JOURNAL  
 MEDLINE  
 21922687  
 PUBMED  
 11925379  
 REFERENCE  
 2 (bases 1 to 3885)  
 Tong, Z.-B., Bondy, C.A., and Nelson, L.M.  
 Direct Submission  
 JOURNAL  
 Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD,  
 NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA  
 FEATURES  
 source  
 1. .3885  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 1. .3885  
 /gene="MATER"  
 1. .3603  
 /gene="MATER"  
 /note="oocyte-specific protein; similar to Mus musculus  
 Mater encoded by GenBank Accession Numbers AF074018 and

CDS  
 gene  
 1. .3603  
 /gene="MATER"  
 /note="oocyte-specific protein; similar to Mus musculus  
 Mater encoded by GenBank Accession Numbers AF074018 and

AF143559-AF143573"  
/codon\_start=1  
/product="maternal-antigen-that-embryos-require protein"  
/protein\_id="AA15549.1"  
/db\_xref="GI:19882273"  
/translation="MKVAGGLELAALILSAPALVTLSGTPCSILPKNPLPQNL  
SSQPCIKMEGDKSLTPSSYGLQWCIYEIDKEEPQTFKELKKKSESTSCIPQETL  
NANVELLALHEYYGASLAWATSIYEENNRLTSEKADMDKRSDEPRLTMR  
QSPKEKVEQISOAVODSATAETKEOETSOAMEOSGATPAETEOEISOAMEOG  
TAETBEQHGSGDPTWDYKSHYMTKFASEDRBSFEMTADAMPQTLAAGFSDRC  
FRPRTVLHGKSGIGKSLARKIYLCMAQGLSYGMFSYFLFVREMRKESVTE  
FISREWPSQAPVTEIMSPERLLFIIDGFDDLSVTLNDKLCMDMAEKQPPTLIR  
SLRKVLPESEFLIVTADYGTKEKLSVSVYLVGSEGRHILHLEERGHEHQ  
KTQGLRAIINRELLDOCVPAVGLICVALQLODVESVAPINQOITGHAALFH  
OLTPRGVRRCLNLEERVYLKRFQMAVEGVMKRSVGGDILVNOGIGSEELALFH  
MNLILPDHSCHEYYTFPHLSIQDFCAALYVLEGLIEPRLCPVYEKTRSMELKA  
GFTHLSLMKRFPLGLVSEDRPLEVLLGCPVLGVQKLLHWLSLQDQPNATTPG  
DTLDAFCLFETQKEFVRLALNSFOEWLPIINQDLIASFCLQHPYRLKRVY  
KGIPEPDESACAVPPLMWRDKTLIEBOWEDPCSMDLTPEHLQDLGSSILTERAM  
KTLCAKLRHPTCKIOTLMFRNAOITPGVQIMRYLMANRNLNSLNGTHLKEDVYM  
ACEALKHFKLLESLDDCGLTACVYKLSOILITTSPLKSLSLAGNKGDOGVTP  
SDALRVSOCLOKLIEBDCGTTAGCGSLAALVNSLTHCLSNLSGNSGNYLNC  
RSKRLPHOSLQRLMNOCHDITAGCGSLAALMNSMLTSLNSNPVEDGVKLCE  
VWRBPSCHLQDLVLVCKLITACCELSVLSRSHLKSIDLTNALGDGVAAFCBG  
LKORNSVLTBLGACGLTSDCBALSLALSNRHLTSLNVONNFSFKMKLCSAF  
ACPTSNLIQILMKMYPVQIRKLEEVQLKPRVIVIDGSWHSFDEDDR YWKN"

## ORIGIN

Query Match 78.5%; Score 3166.2; DB 9; Length 3885;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;  
1 ATGGAAGAGACAATGCTGCTCACCCTTTCCAGCTACGGGGTGCATGCTGCTCATAG 60  
Db 154 ATGGAAGAGACAATGCTGCTCACCCTTTCCAGCTACGGGGTGCATGCTGCTCATAG 213  
Qy 61 CTAGAACAAGAAATTTCAAGCATTTCAAGAAATTTCTAAGAAATTTCTTCAAGATCG 120  
Db 214 CTAGAACAAGAAATTTCAAGCATTTCAAGAAATTTCTAAGAAATTTCTTCAAGATCG 273  
Qy 121 ACCCATGCTCTATTTCCACAGTTTGAATTCAGAAATGCGAACGTGGAATGCTTGGCATC 180  
Db 274 ACCCATGCTCTATTTCCACAGTTTGAATTCAGAAATGCGAACGTGGAATGCTTGGCATC 333  
Qy 181 CTCTTGATGAGTATTTAGAGACATCGTGGCTGGGCTAAGTCAATAGATCTTTGAA 240  
Db 334 CTCTTGATGAGTATTTAGAGACATCGTGGCTGGGCTAAGTCAATAGATCTTTGAA 393  
Qy 241 AACATGAACCTGCGAACCTCTCGGAGAAAGCACGGATGACATGAAAA----- 289  
Db 394 AACATGAACCTGCGAACCTCTCGGAGAAAGCACGGATGACATGAAAAAGCATTCACCA 453  
Qy 290 ----- 289  
Db 454 GAAGATCTCTGAAGCAACGATGACAGCAAGGACCAAGCAAGAAAAAGTGCAGGAATT 513  
Qy 290 -----AAATTCA 297  
Db 514 TCACAAGCTGTGCAACAAGATAGTCCACAGCTGCAGAGACAAAGAAACAGAGAAATTTCA 573  
Qy 298 CAAGCTATGGAACAAGAGGTGCCACAGCAGCAGAGACAGAAAGCAAGAAATTTCAACA 357  
Db 574 CAAGCTATGGAACAAGAGGTGCCACAGCAGAGACAGAAAGCAAGAAATTTCAACA 633  
Qy 358 GCTATGGAACAAGAGGTGCCACAGCAGCAGAGACAGAAAGCAAGAAATTTCAACA 417  
Db 634 GCTATGGAACAAGAGGTGCCACAGCAGCAGAGACAGAAAGCAAGAAATTTCAACA 693  
Qy 418 ACATGGGACTCAAGAGTCAAGTATGACAAATTTGCTGAGAGAGAGATGATGCTGCT 477  
Db 694 ACATGGGACTCAAGAGTCAAGTATGACAAATTTGCTGAGAGAGAGATGATGCTGCT 753  
Qy 478 AGTTTGAAGAACCTGCTGCTGACCTGCGGAATGCAACGTTGGCTGCTGCTTTGAT 537

|||||  
Db 754 AGTTTGAAGAACCTGCTGCTGACCTGCGCGGAATGCAACGTTGGCTGCTGCTTTGAT 813  
Qy 538 TCACAACGGTGGGGCTTCGGGCTCGACAGGTGCTTCGACAGGAAGTACAGAAATTTGGG 597  
Db 814 TCACAACGGTGGGGCTTCGGGCTCGACAGGTGCTTCGACAGGAAGTACAGAAATTTGGG 873  
Qy 598 AAATCGCTCTAGCAGAAAGATCGTCTGCTGGCGCAAGGTGACTCTACCAAGGA 657  
Db 874 AAATCGCTCTAGCAGAAAGATCGTCTGCTGGCGCAAGGTGACTCTACCAAGGA 933  
Qy 658 ATGTTCTCTACGTCCTTCTTCTCCCGTTAAGAGATGACAGCGAAGAAAGAGACACT 717  
Db 934 ATGTTCTCTACGTCCTTCTTCTCCCGTTAAGAGATGACAGCGAAGAAAGAGACAGT 993  
Qy 718 GTCAAGAGTTCATCTCAGAGAGTGGCCAGACTCCAGGCTCCGGTGAACGAGATCATG 777  
Db 994 GTCAAGAGTTCATCTCAGAGAGTGGCCAGACTCCAGGCTCCGGTGAACGAGATCATG 1053  
Qy 778 TCCGACCAAGAAAGCTGTTGTTTCATCATTTGACAGTTCAGTACCTGGCTCTGCTC 837  
Db 1054 TCCGACCAAGAAAGCTGTTGTTTCATCATTTGACAGTTCAGTACCTGGCTCTGCTC 1113  
Qy 838 AACAAATGACAAAGCTCTGCAAAAGACTGAGAGAGCAAGCTCCGTTCACTCTATA 897  
Db 1114 AACAAATGACAAAGCTCTGCAAAAGACTGAGAGAGCAAGCTCCGTTCACTCTATA 1173  
Qy 898 CGCAGTCTGAGAGAGGTCTGCTCCCTGAGTCTTCCGTGACGTCACGTCAGAGAC 957  
Db 1174 CGCAGTCTGAGAGAGGTCTGCTCCCTGAGTCTTCCGTGACGTCACGTCAGAGAC 1233  
Qy 958 GTGGGCAACAGAAAGCTCAAGTCAAGAGTGTCTCCCGTTACTGTTAGTAAAGGA 1017  
Db 1234 GTGGGCAACAGAAAGCTCAAGTCAAGAGTGTCTCCCGTTACTGTTAGTAAAGGA 1293  
Qy 1018 ATCTCCGGGAAACAAAGATTCACCTGCTCTCTTGAAGCGCGGAAATTTGGTACATCAAG 1077  
Db 1294 ATCTCCGGGAAACAAAGATTCACCTGCTCTCTTGAAGCGCGGAAATTTGGTACATCAAG 1353  
Qy 1078 ACAAGAGGTGGCGTGGATCATGAAACAACGTGAGTGTCTCCGACAGTGCAGGTGCC 1137  
Db 1354 ACAAGAGGTGGCGTGGATCATGAAACAACGTGAGTGTCTCCGACAGTGCAGGTGCC 1413  
Qy 1138 GCCGTGGGCTCTCATCTGCGTGGCTCGTCAAGTCAAGACGTGTGGGAGAGAGGCTC 1197  
Db 1414 GCCGTGGGCTCTCATCTGCGTGGCTCGTCAAGTCAAGACGTGTGGGAGAGAGGCTC 1473  
Qy 1198 GCCCTTTCAACAAAGCTTCACAGGCTTCAGCGCTTTTGTGTTCAATCACTCAACC 1257  
Db 1474 GCCCTTTCAACAAAGCTTCACAGGCTTCAGCGCTTTTGTGTTCAATCACTCAACC 1533  
Qy 1258 CCTGGAAGCGGTGTCGGGCGGTGCTCAATCTGGAAGAAAGTGTCTGGAAGCGCTTC 1317  
Db 1534 CCTGGAAGCGGTGTCGGGCGGTGCTCAATCTGGAAGAAAGTGTCTGGAAGCGCTTC 1593  
Qy 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGTTGATGTTGAGAGAGCTC 1377  
Db 1594 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGTTGATGTTGAGAGAGCTC 1653  
Qy 1378 ATGCTTCAAGACTCGGAGAGTCTGAGCTCCGTGCTGTTTCAATGAAATCACTCTTC 1437  
Db 1654 ATGCTTCAAGACTCGGAGAGTCTGAGCTCCGTGCTGTTTCAATGAAATCACTCTTC 1713  
Qy 1438 CCAAGACAGCACTGTGAGAGTACTAACACTTCTTCAACCTCAAGTCCAGGACCTTCGT 1497  
Db 1714 CCAAGACAGCACTGTGAGAGTACTAACACTTCTTCAACCTCAAGTCCAGGACCTTCGT 1773  
Qy 1498 GCCGCTTTGACTAGTGTGAGAGGCTGGAATCGAGCAGGCTCTGCGCTCTGATC 1557  
Db 1774 GCCGCTTTGACTAGTGTGAGAGGCTGGAATCGAGCAGGCTCTGCGCTCTGATC 1833  
Qy 1558 GTTGAAGAACAAAGAGTTCATGAGCTTAAACAGGCAAGCTTCAATATCACTGCTT 1617

Db 1834 GTTGAGAGACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCATTCACCTGCTT 1893  
Qy 1618 TGGATGAGAGCTTTCTGTTTGGCCCTGAGAGGAAAGATGAGAGGCACTGAGAGTCC 1677  
Db 1894 TGGATGAGAGCTTTCTGTTTGGCCCTGAGAGGAAAGATGAGAGGCACTGAGAGTCC 1953  
Qy 1678 CTGCTGGGCTGCTCCGTTTCCCTGGGGGTGAGACAGAACTTCTGCACTGGGCTCTCTG 1737  
Db 1954 CTGCTGGGCTGCTCCGTTTCCCTGGGGGTGAGAGGAAAGCTTCTGCACTGGGCTCTCTG 2013  
Qy 1738 TTGGGTCAGAGCCTTAATGCAACCCAGAGAGACCTCTGAGCGCTTTCACCTGCTT 1797  
Db 2014 TTGGGTCAGAGCCTTAATGCAACCCAGAGAGACCTCTGAGCGCTTTCACCTGCTT 2073  
Qy 1798 TTGAGACTCAAGACAAAGATTTGTTGGCTTGGCATTAACAGCTTCCAAAGATGTG 1857  
Db 2074 TTGAGACTCAAGACAAAGATTTGTTGGCTTGGCATTAACAGCTTCCAAAGATGTG 2133  
Qy 1858 CTTCGCAATTAACAGACCTGAGCTTGAATGATCTTCTTGGCTCCAGAGCTGCTG 1917  
Db 2134 CTTCGCAATTAACAGACCTGAGCTTGAATGATCTTCTTGGCTCCAGAGCTGCTG 2193  
Qy 1918 TATTTGCGGAAATTCGGGTGAGATGCAAAAGGATCTTCCAAAGATGAGTCCGCTGAG 1977  
Db 2194 TATTTGCGGAAATTCGGGTGAGATGCAAAAGGATCTTCCAAAGATGAGTCCGCTGAG 2253  
Qy 1978 GATATGCTGTTGCTCTCTTAATGATGCGGGATTAAGACCTTGAAGAGAGAGTGGAA 2037  
Db 2254 GATATGCTGTTGCTCTCTTAATGATGCGGGATTAAGACCTTGAAGAGAGAGTGGAA 2313  
Qy 2038 GATTTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGAGACCTGGGAGAGAC 2097  
Db 2314 GATTTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGAGACCTGGGAGAGAC 2373  
Qy 2098 ATCTGAGACAGAGGAGGCAATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAG 2157  
Db 2374 ATCTGAGACAGAGGAGGCAATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAG 2433  
Qy 2158 ATACGAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGTGTGAGAGACTCTGAGAG 2217  
Db 2434 ATACGAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGTGTGAGAGACTCTGAGAG 2493  
Qy 2218 ATGCTATGAGGCAACCTGAACCTTAAGATCCCTCACTTGGAGAGGACCACTGAAGAA 2277  
Db 2494 ATGCTATGAGGCAACCTGAACCTTAAGATCCCTCACTTGGAGAGGACCACTGAAGAA 2553  
Qy 2278 GAGGATGTAAGAGTGGCTGTGAAGCTTAAACCAACCAATGTTTGTGAGTCTTGG 2337  
Db 2554 GAGGATGTAAGAGTGGCTGTGAAGCTTAAACCAACCAATGTTTGTGAGTCTTGG 2613  
Qy 2338 AGGCTGATGCTGTGATGATGAACCAATGCTTGAACCTGAAGATCTCCAAATCTTACG 2397  
Db 2614 AGGCTGATGCTGTGATGATGAACCAATGCTTGAACCTGAAGATCTCCAAATCTTACG 2673  
Qy 2398 ACCCTCCGAGCTGAATCTGTGAGCTTGGAGGAAACAAAGGTACAGACCAAGGAGATA 2457  
Db 2674 ACCCTCCGAGCTGAATCTGTGAGCTTGGAGGAAACAAAGGTACAGACCAAGGAGATA 2733  
Qy 2458 ATGCTCTCAATGATGCTTGAAGATCTCCAGTGGGCTCTGAGAGCTGATCTGAG 2517  
Db 2734 ATGCTCTCAATGATGCTTGAAGATCTCCAGTGGGCTCTGAGAGCTGATCTGAG 2793  
Qy 2518 GACTGTGATCAACAGCAACGAGTGGCTGAGAGTCTGAGCTTGGCTGAGAGCTGAG 2577  
Db 2794 GACTGTGATCAACAGCAACGAGTGGCTGAGAGTCTGAGCTTGGCTGAGAGCTGAG 2853  
Qy 2578 AGCTTGAACACCTGAGCTATCAACAGCTGGGGAACAGAGGTAAATCTTACG 2637  
Db 2854 AGCTTGAACACCTGAGCTATCAACAGCTGGGGAACAGAGGTAAATCTTACG 2913  
Qy 2638 TGTGATCATGAGGCTTCCCACTGTAGTCTGAGAGGCTGATCAATCATGAGCAC 2697  
Db 2914 TGTGATCATGAGGCTTCCCACTGTAGTCTGAGAGGCTGATCAATCATGAGCAC 2973

Qy 2698 CTGAGACAGGCTGAGCTGTTTCTTCACTTGGCTTAATGGGTAACTAGGCTGACG 2757  
Db 2974 CTGAGACAGGCTGAGCTGTTTCTTCACTTGGCTTAATGGGTAACTAGGCTGACG 3033  
Qy 2758 CACTGAGCTTGAATGAACCTGTGAGAGCAATGAGCTGAAGCTTGTGAGAGTCC 2817  
Db 3034 CACTGAGCTTGAATGAACCTGTGAGAGCAATGAGCTGAAGCTTGTGAGAGTCC 3093  
Qy 2818 ATGAGAGAACCAATCTTGAATCTTCAAGACCTGAGATTGTAAGTGTATCTCACCGCC 2877  
Db 3094 ATGAGAGAACCAATCTTGAATCTTCAAGACCTGAGATTGTAAGTGTATCTCACCGCC 3153  
Qy 2878 GGTGCTGTGAGAGCTGCTGCTGTGATCTGAGAGACACACCTGAAGAGCTGTGAT 2937  
Db 3154 GGTGCTGTGAGAGCTGCTGCTGTGATCTGAGAGACACACCTGAAGAGCTGTGAT 3213  
Qy 2938 CTACAGGACCAATGAGCTGAGTGAAGCTGGGCTGAGCTGTGAGAGGAGCTGAAGCA 2997  
Db 3214 CTACAGGACCAATGAGCTGAGTGAAGCTGGGCTGAGCTGTGAGAGGAGCTGAAGCA 3273  
Qy 2998 AAGAACAGTGTCTGAGAGACTGGGCTTGAAGAGCTGATGATCTTGTGATGCTGT 3057  
Db 3274 AAGAACAGTGTCTGAGAGACTGGGCTTGAAGAGCTGATGATCTTGTGATGCTGT 3333  
Qy 3058 GAGGACCTCTGCTGGGCTTCTGCAACCGGATCTGACCAAGCTTAAACCTGTGAG 3117  
Db 3334 GAGGACCTCTGCTGGGCTTCTGCAACCGGATCTGACCAAGCTTAAACCTGTGAG 3393  
Qy 3118 AATAACTTCAGTCCCAAGAGATGAAGCTGTGAGGCTTGTGCTGTCCACGCTCT 3177  
Db 3394 AATAACTTCAGTCCCAAGAGATGAAGCTGTGAGGCTTGTGCTGTCCACGCTCT 3453  
Qy 3178 AACTTACAGATTAATGAGCTGTGAGAAATGAGAGTACCTGTGCAATTAAGAGAGTGTG 3237  
Db 3454 AACTTACAGATTAATGAGCTGTGAGAAATGAGAGTACCTGTGCAATTAAGAGAGTGTG 3513  
Qy 3238 GAGGAGTGTGAGCTTAACCTGAGTGTGATGAGTGTGAGTGTGATCTTGTGAT 3297  
Db 3514 GAGGAGTGTGAGCTTAACCTGAGTGTGATGAGTGTGAGTGTGATCTTGTGAT 3573  
Qy 3298 GAGGATGAGCGACAC 3312  
Db 3574 GAGGATGAGCGGTAC 3588

RESULT 3  
AX427610  
LOCUS  
DEFINITION Sequence 23 from Patent WO0232955.  
ACCESSION AX427610  
VERSION AX427610.1 GI:21537730  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE 1  
AUTHORS Nelson,L.M. and Tong,Z.-B.  
TITLE Human gene critical to fertility  
JOURNAL Patent: WO 0232955-A 23 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
FEATURES  
source  
1..3900  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..3603  
/note="unassigned protein product"  
/codon\_start=1  
/protein\_id="CAD35973.1"  
/db\_xref="GI:21537731"  
/translation="MKVAGSLLEGAALLSAPRALVTLSTGPTCSILPKNPILFPQNL



SSOPCIMEGDKSLTFSYGLMCLYEIDKEEPOTFKELLKKSSSESTCS1POFEIE  
NANVECLALTHEYYGASLAWATSI1FENNRLTLEKADDDKRSBEDPALTMD  
OGPEKREVPGISOAVODSATAETKOEISOAMEOBGATLEBEDEISOBOGA  
TAATBEQGHGSDTMDYKSHMTKPAEBEDRBSFENTADMPMOTLAGPDSRNG  
FRPTVLDHGSIGKSLARIVLCAHOGSLYOGMPSYFLEKREORKESSVIE  
FISEWPDSDQAPTEIMSRPERLFIIDGFDDLSVANDKICKDVAEKQPPETLIR  
SLAKVILPESFLIVTVADVOTEKIKESVSPRYLVGHSGBORIHLLERGBEQ  
KTQGLRAIINRRELDQOVAVGSLICVALQLODVESVAPENQOTLEIGDGVALCEG  
OLTRGVRRCLNLEERVYLKRFORMAVEGWNRKSVEDGDLMLGQSESEMLKQ  
MNILIPDSHCEVYTPFLSLDPCALVYVLBGLLEIPLACPLVYKRTSMELKQ  
GFTHSLMKRFLFLVSEDVRPLEVILGCPVPLGYKOKILHVSILGOOPNATTTG  
DTLDAFCLFETQDEPVRLALNSFOEWLPEINLNDLASSFCLQKDEKPYRKLIRVY  
KGPREDSEAECAVPEPLMMDKTLIEBOWDFCSMLGTHPHLEOLDGSSILIERAM  
KITLCAKRLHPCKIOTLMFRNAOTTPGVOLHRIYMANRNLINLGGTILKEEDVEM  
ACELAKHPKCLLESRLDDCGCTHACTYKISLITLTPSLKSLIAGNKVTDGVTPL  
SDALRVOCALQKILIEDCGTATGCGSLALVSNSTIHLCSNNSISNGEYNILC  
RSKRLPHOSLORMLNOCHLDTAGCSLALMLKMSWITLHSLSNPVEDVDGVALCE  
VMRPSCHLODELKCHLTAACESLSVLSRSHLKLIDLTDNALGDGVALCEG  
LKQNSVYTRIGLAVCGITSDCEBALSTALSNLTVONNFPKGMKLCGSAF  
ACPTSNLQIIGLMKQYPVQIRKLIEVOLLKPRVLDGSHWSFDEDDRVMKKN"

## ORIGIN

Query Match 78.5%; Score 3166.2; DB 6; Length 3900;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

OY 1 ATGGAAGGAGCAATATGCTACCTTTCCAGCTACGGGCTGCATGGTGTCTATGAG 60  
DB 154 ATGGAAGGAGCAATATGCTACCTTTCCAGCTACGGGCTGCATGGTGTCTATGAG 213  
OY 61 CTAGAAGAAGAAATTTCAAGCATTTCAAGAAATTTCAAGAAATTTCTTCAAGATCG 120  
DB 214 CTAGAAGAAGAAATTTCAAGCATTTCAAGAAATTTCAAGAAATTTCTTCAAGATCG 273  
OY 121 ACCAATGCTCTATTTCCACAGTTTGAATTCGAAGATTCGAAGATTTGCTGGCACTC 180  
DB 274 ACCAATGCTCTATTTCCACAGTTTGAATTCGAAGATTCGAAGATTTGCTGGCACTC 333  
OY 181 CTCTTGATGATTTATGAGCATCGTGGCTGGGCTGATGATGATGATGATGATGATGAT 240  
DB 334 CTCTTGATGATTTATGAGCATCGTGGCTGGGCTGATGATGATGATGATGATGATGAT 393  
OY 241 AACATGAACCTGCGAACCCTCTCGGAGAAAGGACGGGATGACATGAAAA----- 289  
DB 394 AACATGAACCTGCGAACCCTCTCGGAGAAAGGACGGGATGACATGAAAAAGCATTCACCA 453  
OY 290 ----- 289  
DB 454 GAAGATCTGAAGCAACGATGATGACCAAGGACCAAGCAAGAAAGTCCAGGAATTT 513  
OY 290 -----AATTTCA 297  
DB 514 TCACAACTGTGCAACAAGATGAGCCACAGCTGACAGACAAAGAAAGGAAATTTCA 573  
OY 298 CAAGCTATGAACAAGAGGTGCGACACAGACAGACAGACAGAAATTTTCAACA 357  
DB 574 CAAGCTATGAACAAGAGGTGCGACACAGACAGACAGAAAGAAATTTTCAACA 633  
OY 358 GCTATGAACAAGAGGTGCGACACAGACAGACAGACAGAAACAAGACATGAGGTGAC 417  
DB 634 GCTATGAACAAGAGGTGCGACACAGACAGACAGAAACAAGACATGAGGTGAC 693  
OY 418 ACATGGGACTCAAGAGTCAAGTGAAGCAAAATTCGTGAGGAGGAGAGTACGTGCT 477  
DB 694 ACATGGGACTCAAGAGTCAAGTGAAGCAAAATTCGTGAGGAGGAGAGTACGTGCT 753  
OY 478 AGTTTGAACAACATGCTGCTGACTGCGCGGAAATGCAAAAGTGGCTGTGCTTTGAT 537  
DB 754 AGTTTGAACAACATGCTGCTGACTGCGCGGAAATGCAAAAGTGGCTGTGCTTTGAT 813  
OY 538 TCAAGCCGGTGGGCTTCCGGCTTGCACGGTGGTCTTGCACCGAAAGTCAAGAAATTTGGG 597  
DB 814 TCAAGCCGGTGGGCTTCCGGCTTGCACGGTGGTCTTGCACCGAAAGTCAAGAAATTTGGG 873

OY 598 AATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGACTCTACAGAGA 657  
DB 874 AATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGGTGACTCTACAGAGA 933  
OY 658 ATGTTCTCTACGCTCTTCTCTCCCGTTAGAGATGCAAGCGGAGAAAGAGACAGT 717  
DB 934 ATGTTCTCTACGCTCTTCTCTCCCGTTAGAGATGCAAGCGGAGAAAGAGACAGT 993  
OY 718 GTCAAGATTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGGAGAGGAGATCATG 777  
DB 994 GTCAAGATTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGGAGAGGAGATCATG 1053  
OY 778 TCCCGACAGAAAGGCTGTTGTTTATGATGACGTTTTCATGACCTGGCTGTGCTC 837  
DB 1054 TCCCGACAGAAAGGCTGTTGTTTATGATGACGTTTTCATGACCTGGCTGTGCTC 1113  
OY 838 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACGCTCCGTTCACTCATYA 897  
DB 1114 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACGCTCCGTTCACTCATYA 1173  
OY 898 CGCAGTCTGCTGAGAGAGTCTGCTCCGAGAGTCTTCTGATCGTCAAGCTGAGAGAC 957  
DB 1174 CGCAGTCTGCTGAGAGAGTCTGCTCCGAGAGTCTTCTGATCGTCAAGCTGAGAGAC 1233  
OY 958 GTGGGACAGAGAAAGCTCAAGTCAAGAGTGTGCTCCCGTTACCTGTTAGTTAGAGA 1017  
DB 1234 GTGGGACAGAGAAAGCTCAAGTCAAGAGTGTGCTCCCGTTACCTGTTAGTTAGAGA 1293  
OY 1018 ATCTCCGGGGAACAAAGATTCATCTGCTCTTGAAGCGCGGAGATGTTGAGATCAAG 1077  
DB 1294 ATCTCCGGGGAACAAAGATTCATCTGCTCTTGAAGCGCGGAGATGTTGAGATCAAG 1353  
OY 1078 AACAAAGGTTGCGTGGATCAAGAAACCGTGAAGTGTCTGACAGAGTGCAGAGTCCC 1137  
DB 1354 AACAAAGGTTGCGTGGATCAAGAAACCGTGAAGTGTCTGACAGAGTGCAGAGTCCC 1413  
OY 1138 GCCGTGGGCTCTCTCATCTGAGTGGCCCTGACGTGACAGACGTTGAGGAGAGAGCTC 1197  
DB 1414 GCCGTGGGCTCTCTCATCTGAGTGGCCCTGACGTGACAGACGTTGAGGAGAGAGCTC 1473  
OY 1198 GCCCCTTTCACCAAGCTCTACAGGCTTGACAGCGCGCTTTGTTGTTTCAATCACTACC 1257  
DB 1474 GCCCCTTTCACCAAGCTCTACAGGCTTGACAGCGCGCTTTGTTGTTTCAATCACTACC 1533  
OY 1258 CCTGAGGCGTGGTCCGCGCTGTCTCAATCTGAGGAGAAAGATGTCCTGAAGCGCTTC 1317  
DB 1534 CCTGAGGCGTGGTCCGCGCTGTCTCAATCTGAGGAGAAAGATGTCCTGAAGCGCTTC 1593  
OY 1318 TGCCGTATGCTGTGAGAGGAGTGTGAATAGAAAGTCAAGTGTGTTGATGTTGACGACTC 1377  
DB 1594 TGCCGTATGCTGTGAGAGGAGTGTGAATAGAAAGTCAAGTGTGTTGATGTTGACGACTC 1653  
OY 1378 ATGTTCAAGAGACTCGGGAGTGTGAGCTCCGTGCTGTTTCAATGAACATCTTCTC 1437  
DB 1554 ATGTTCAAGAGACTCGGGAGTGTGAGCTCCGTGCTGTTTCAATGAACATCTTCTC 1713  
OY 1438 CCAAGACGACCTGTGAGAGTACTAACAATCTTCCACCTCACTCACTCACTCACTTCTGT 1497  
DB 1714 CCAAGACGACCTGTGAGAGTACTAACAATCTTCCACCTCACTCACTCACTCACTTCTGT 1773  
OY 1498 GCCGCTTGTACTAGTGTGTTAGAGGCTGGAATTCAGACCAAGCTCTGCGCTCTGATC 1557  
DB 1774 GCCGCTTGTACTAGTGTGTTAGAGGCTGGAATTCAGACCAAGCTCTGCGCTCTGATC 1833  
OY 1558 GTTGAAGAACAAAGAGTCAATGAGCTTAAACAGGCAAGCTTCAATATCACTGCTT 1617  
DB 1834 GTTGAAGAACAAAGAGTCAATGAGCTTAAACAGGCAAGCTTCAATATCACTGCTT 1893  
OY 1818 TGAATGAAGGTTTCTTGTGCTGTGAGGAGAAAGCTGAAGAGGCACTGGAAGTGC 1677  
DB 1894 TGAATGAAGGTTTCTTGTGCTGTGAGGAGAAAGCTGAAGAGGCACTGGAAGTGC 1953



```

Qy 1678 CTGCTGGAGCTGTCCCGTCCCTGGGGGTGAAGCAGAACCTTGTGCACTGGGTCTCTCTG 1737
Db 1954 CTGCTGGAGCTGTCCCGTCCCTGGGGGTGAAGCAGAACCTTGTGCACTGGGTCTCTCTG 2013
Qy 1738 TTGGGTGACACCTTAAATGCCACCAACCCAGAGAGACCTTGGAGCGCTTCCACTGTCTT 1797
Db 2014 TTGGGTGACACCTTAAATGCCACCAACCCAGAGAGACCTTGGAGCGCTTCCACTGTCTT 2073
Qy 1798 TTGAGACTCAAGACAAAGATTGTTGGCTTGGCACTTAAACAGCTTCCAGAGAGTGGG 1857
Db 2074 TTGAGACTCAAGACAAAGATTGTTGGCTTGGCACTTAAACAGCTTCCAGAGAGTGGG 2133
Qy 1858 CTTCGAGTTAACAGAACCTTGAATCTTGTCTTCCCTTCCCTCCAGCACTGTCCG 1917
Db 2134 CTTCGAGTTAACAGAACCTTGAATCTTGTCTTCCCTTCCCTCCAGCACTGTCCG 2193
Qy 1918 TATTTGGGAAAAATTCGGGTGATGTCAAGGGATTTTCCCAAGAGATGATCCGTGAG 1977
Db 2194 TATTTGGGAAAAATTCGGGTGATGTCAAGGGATTTTCCCAAGAGATGATCCGTGAG 2253
Qy 1978 GCATGTCCTGTCCTCTATGAGTGGGGATTAAGACCTTCAATTGAGAGAGTGGGA 2037
Db 2254 GCATGTCCTGTCCTCTATGAGTGGGGATTAAGACCTTCAATTGAGAGAGTGGGA 2313
Qy 2038 GATTTCTGCTCATGCTTGGACACCACTGCGGACCTGGAACCTGGGACAGCAC 2097
Db 2314 GATTTCTGCTCATGCTTGGACACCACTGCGGACCTGGAACCTGGGACAGCAC 2373
Qy 2098 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAG 2157
Db 2374 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAG 2433
Qy 2158 ATACGACCCCTGATGTTTGAATGACAGATTAACCTGCTGTGTGAGAGACTCTGAGAG 2217
Db 2434 ATACGACCCCTGATGTTTGAATGACAGATTAACCTGCTGTGTGAGAGACTCTGAGAG 2493
Qy 2218 ATCTGATGAGCCCAACCTTAAGATCCCTCAACTTGGGAGGACCCACCTGAAGGA 2277
Db 2494 ATCTGATGAGCCCAACCTTAAGATCCCTCAACTTGGGAGGACCCACCTGAAGGA 2553
Qy 2278 GAGGATGTAAGATGGCGTGTGAGACCTTAAACACCCAAATGTTTGTGAGTCTTGG 2337
Db 2554 GAGGATGTAAGATGGCGTGTGAGACCTTAAACACCCAAATGTTTGTGAGTCTTGG 2613
Qy 2338 AGGCTGATGCTGTGTGATGACCACTGCTTAACTGAGATCTCCAAATCTTACG 2397
Db 2614 AGGCTGATGCTGTGTGATGACCACTGCTTAACTGAGATCTCCAAATCTTACG 2673
Qy 2398 ACCTCCCGCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAAGGTGACAGACAGGAGTA 2457
Db 2674 ACCTCCCGCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAAGGTGACAGACAGGAGTA 2733
Qy 2458 ATGCTCTCACTGATGCTTGAAGTCTCCAGTCCGCTGACAGAGCTGATCTGAG 2517
Db 2734 AAGCTCTCACTGATGCTTGAAGTCTCCAGTCCGCTGACAGAGCTGATCTGAG 2793
Qy 2518 GACTGTGACATCAACCAACCGGTTGCCAGAGTCTGGCCTCAGCCCTGTGACAAACCGG 2577
Db 2794 GACTGTGACATCAACCAACCGGTTGCCAGAGTCTGGCCTCAGCCCTGTGACAAACCGG 2853
Qy 2578 AGCTTGACACACCTGTGCTTATCCAAACACGCTGGGGAACGAGGTGTAATCTACTG 2637
Db 2854 AGCTTGACACACCTGTGCTTATCCAAACACGCTGGGGAACGAGGTGTAATCTACTG 2913
Qy 2638 TGTGATCACTAGAGCTTCCCACTGTAGTCTGACAGAGCTGATCTGTAATCAGTCCAG 2697
Db 2914 TGTGATCACTAGAGCTTCCCACTGTAGTCTGACAGAGCTGATCTGTAATCAGTCCAG 2973
Qy 2698 CTGACACGAGCTGCTGTGTTTCTTGCACCTTGGCTTATGAGTACTCATGCTGAG 2757
Db 2974 CTGACACGAGCTGCTGTGTTTCTTGCACCTTGGCTTATGAGTACTCATGCTGAG 3033
Qy 2758 CACCTGACCTTGAATGAACCTGTGTGAAGACATGCGTGAAGCTTCTGTGAGGTC 2817

```

```

Db 3034 CACTGAGCCTTAGATGAACCTGTGGAAGCAATGGGTGAACCTTGTGCGCAGGTC 3093
Qy 2818 ATGAGAGAACCATCTTGTATCTCCAGACCTGAGAGTTGTGTAATGTATCTCACCGCC 2877
Db 3094 ATGAGAGAACCATCTTGTATCTCCAGACCTGAGAGTTGTGTAATGTATCTCACCGCC 3153
Qy 2878 GGTGCTGTGAGAGCTGTGCTGTGTGATCTGAGAGAGACACCTGAAGAGCGCTGAT 2937
Db 3154 GGTGCTGTGAGAGCTGTGCTGTGTGATCTGAGAGAGACACCTGAAGAGCGCTGAT 3213
Qy 2938 CTACAGGACATGCTGCTGTGAGACCGTGGGTGCTGCGCTGTGAGAGGACTGAAGCAA 2997
Db 3214 CTACAGGACATGCTGCTGTGAGACCGTGGGTGCTGCGCTGTGAGAGGACTGAAGCAA 3273
Qy 2998 AAGAACAGTGTCTGACAGAGACTCGGCTTGAAGGACATGTGACTGACTTGTGATGCTGT 3057
Db 3274 AAGAACAGTGTCTGACAGAGACTCGGCTTGAAGGACATGTGACTGACTTGTGATGCTGT 3333
Qy 3058 GAGGACCTCTGCTGCTGCTTCTGCAACCGGACATCTGAACAGCTAAACCTGGTGCAG 3117
Db 3334 GAGGACCTCTGCTGCTGCTTCTGCAACCGGACATCTGAACAGCTAAACCTGGTGCAG 3393
Qy 3118 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGTGCGCTTGGCTGTCCACGTCT 3177
Db 3394 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGTGCGCTTGGCTGTCCACGTCT 3453
Qy 3178 AACTTACAGATTAATGGCTGTGAAATGGACATGACCTGTGCAATAAGAGAGTGTCTG 3237
Db 3454 AACTTACAGATTAATGGCTGTGAAATGGACATGACCTGTGCAATAAGAGAGTGTCTG 3513
Qy 3238 GAGGAGTGCAGCTACTCAAGCCCGGAGTCCGTAAATTGACGTAAGTGGACTTTTGTAT 3297
Db 3514 GAGGAGTGCAGCTACTCAAGCCCGGAGTCCGTAAATTGACGTAAGTGGACTTTTGTAT 3573
Qy 3298 GAAGATGACCGGACAC 3312
Db 3574 GAAGATGACCGGATAC 3588

```

```

RESULT 4
AX704823 3830 bp DNA linear PAT 04-APR-2003
LOCUS AX704823
DEFINITION Sequence 3 from Patent EP1285964.
ACCESSION AX704823
VERSION AX704823.1 GI:29561488
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Wells, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.
Human mifer proteins
Patent: EP 1285964-A 3 26-FEB-2003;
JOURNAL SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES
source
1..3830
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

```

ORIGIN
Query Match 77.8%; Score 3139; DB 6; Length 3830;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 3282; Conservative 0; Mismatches 30; Indels 105; Gaps 2;
Qy 1 ATGAGAGAGCAAAATGCTACCTTTCCAGCTACGGGCTGCAATGTGTCTATGAG 60
Db 1 ATGAGAGAGCAAAATGCTACCTTTCCAGCTACGGGCTGCAATGTGTCTATGAG 60
Qy 61 CTAGACAAAGAAATTTCAACATTTCAAGAAATTTACTAAAGAAATCTTCGAATCG 120

```

D6 61 CTAGACAGAGAAATTTACACATTCAGAGAAATTACTAAAGAAATCTTCAGAAATCG 120  
QY 121 ACCACATGCTCTATTCCACAGTTTGAAATCGAGATGCGCAAGTGGAAATGTCGGCACTC 180  
Db 121 ACCACATGCTCTATTCCACAGTTTGAAATCGAGATGCGCAAGTGGAAATGTCGGCACTC 180  
QY 181 CTTTTCATGAGATATTTAGAGATCGCTGCGCTGGCTAGCGTCCATTTAGCATCTTTGAA 240  
Db 181 CTTTTCATGAGATATTTAGAGATCGCTGCGCTGGCTAGCGTCCATTTAGCATCTTTGAA 240  
QY 241 AACATGAACCTGCGAACCTCTCGGAGAAAGGACGGGATGACATGAAAAAATTTACAA 300  
Db 241 AACATGAACCTGCGAACCTCTCGGAGAAAGGACGGGATGACATGAAAAAATTTACAA 300  
QY 301 GCT-----ATGGAACAAGAGTGCACAGACGACAGACAGAGAACAA- 344  
Db 301 GATCCTGAGACAGATGACTGACCCAGACCAAGCAAGAAAAATGTCAGAAAAATPAA 360  
QY 345 ----- 344  
Db 361 TATGCAATGACTAATGCTTATCTTGGGGGTGCTGACATCTGCACTGCAATATPAAACAC 420  
QY 345 -----AATAATTTCAAGCTATGGAACAAGAGT 375  
Db 421 AAGTATGTTGAATTCATCTTCTTTTGCAAGAAATTTCAAGCTATGGAACAAGAGT 480  
QY 376 GCGACAGCAGAGACAGAGAAACAAGACATGGAGGTGACATGGGACTTAACAAGT 435  
Db 481 GCGACAGCAGAGACAGAGAAACAAGACATGGAGGTGACATGGGACTTAACAAGT 540  
QY 436 CACGTGATGACCAATTCGTGAGAGAGAGATGACGTGATGTTTGAACAACATGCT 495  
Db 541 CACGTGATGACCAATTCGTGAGAGAGAGATGACGTGATGTTTGAACAACATGCT 600  
QY 496 GCTGATCGCGCGGAATGCAAACTGGCTGTGCTTTGATTCAGACCGGTGGGCTTC 555  
Db 601 GCTGATCGCGCGGAATGCAAACTGGCTGTGCTTTGATTCAGACCGGTGGGCTTC 660  
QY 556 CGGCGCTGCGACGAGGTCTGCAACGAAAGTCAAGAAATGGGAAATCGGCTCTAGCCCA 615  
Db 661 CGGCGCTGCGACGAGGTCTGCAACGAAAGTCAAGAAATGGGAAATCGGCTCTAGCCCA 720  
QY 616 AGGATCGTGTGTGCTGGCGCAAGTGAAGCTTAACAGGAAATGTTCTCTCAAGTCTTC 675  
Db 721 AGGATCGTGTGTGCTGGCGCAAGTGAAGCTTAACAGGAAATGTTCTCTCAAGTCTTC 780  
QY 676 TTCTCTCCCTTTAGAGATGACGCGGAAGAGAGACAGTGTCAAGATTCATCTCC 735  
Db 781 TTCTCTCCCTTTAGAGATGACGCGGAAGAGAGACAGTGTCAAGATTCATCTCC 840  
QY 736 AGGAGTGGCCAGATCCCAAGGCTCCGCTGAACGAGATCATGTCCCGACCAAGAAAGCTG 795  
Db 841 AGGAGTGGCCAGATCCCAAGGCTCCGCTGAACGAGATCATGTCCCGACCAAGAAAGCTG 900  
QY 796 TTGTTTCATCATTTGACGATTTGATGACCTGGGCTCTGCTCAACAATGACACAAGCTC 855  
Db 901 TTGTTTCATCATTTGACGATTTGATGACCTGGGCTCTGCTCAACAATGACACAAGCTC 960  
QY 856 TGCAGAGCTGGGCTGAGAGACGCTCTGTTCAACCTCTCATACGAGTCTGTGAGAG 915  
Db 961 TGCAGAGCTGGGCTGAGAGACGCTCTGTTCAACCTCTCATACGAGTCTGTGAGAG 1020  
QY 916 GTCCTGCTCCGAGTCTCTCTGATGATGACCCGTCAGAGACGTTGGGACAGAGAGCTC 975  
Db 1021 GTCCTGCTCCGAGTCTCTCTGATGATGACCCGTCAGAGACGTTGGGACAGAGAGCTC 1080  
QY 976 AAGTCAGAGTCTGTCTCTCCCGTTACCTGTTAGTAGAGATCTCCGGGAAACAAAGA 1035  
Db 1081 AAGTCAGAGTCTGTCTCTCCCGTTACCTGTTAGTAGAGATCTCCGGGAAACAAAGA 1140  
QY 1036 ATCCACTTGTCTCTTGAAGCGGGGATTTGTGAGCATCAAGAGACCAAGGGTGTGCTCG 1095  
Db 1141 ATCCACTTGTCTCTTGAAGCGGGGATTTGTGAGCATCAAGAGACCAAGGGTGTGCTCG 1200

QY 1096 ATCATGAACAACCGTAGCTGTGACAGATGCGCAGGTGCGCGGTGGCTCTCATC 1155  
Db 1201 ATCATGAACAACCGTAGCTGTGACAGATGCGCAGGTGCGCGGTGGCTCTCATC 1260  
QY 1156 TGCCTGCGCTGCACTGTCAGAGACGTGTGGGGAGAGCGTCCGCCCTTCAACCAAG 1215  
Db 1261 TGCCTGCGCTGCACTGTCAGAGACGTGTGGGGAGAGCGTCCGCCCTTCAACCAAG 1320  
QY 1216 CTCACAGGCGTCACGCGCTTTGTGTTCATCAGTCAACCCCTGAGGCGTGTCCG 1275  
Db 1321 CTCACAGGCGTCACGCGCTTTGTGTTCATCAGTCAACCCCTGAGGCGTGTCCG 1380  
QY 1276 CGCTGTCTCAATCTGAGAAAGATGTCTGAAAGCGTTCGTGCGTATGTGCTGAG 1335  
Db 1381 CGCTGTCTCAATCTGAGAAAGATGTCTGAAAGCGTTCGTGCGTATGTGCTGAG 1440  
QY 1336 GAGTGTGGAATAGAGAGTCAGTGTTGATGTGACGACCTCATGTTCAAGACTCGG 1395  
Db 1441 GAGTGTGGAATAGAGAGTCAGTGTTGATGTGACGACCTCATGTTCAAGACTCGG 1500  
QY 1396 GAGTGTGAGCTCGTGTCTGTTTCAATGAACAATCCTTCCACAGCAGCTGTGAG 1455  
Db 1501 GAGTGTGAGCTCGTGTCTGTTTCAATGAACAATCCTTCCACAGCAGCTGTGAG 1560  
QY 1456 GAGTACTACACCTTCTTCCACTCACTCAAGTCTTGTGCGCGCTTGTACTACGTG 1515  
Db 1561 GAGTACTACACCTTCTTCCACTCACTCAAGTCTTGTGCGCGCTTGTACTACGTG 1620  
QY 1516 TTAGAGGCGCTGGAATGACAGCAGCTCTGCGCTTGTACCTTTGAGAAAGCAAAAG 1575  
Db 1621 TTAGAGGCGCTGGAATGACAGCAGCTCTGCGCTTGTACCTTTGAGAAAGCAAAAG 1680  
QY 1576 TCCATGAGGCTTAAACAGGAGGCTTCATATCCACTCGCTTGGAGTGAAGCGTTCTTG 1635  
Db 1681 TCCATGAGGCTTAAACAGGAGGCTTCATATCCACTCGCTTGGAGTGAAGCGTTCTTG 1740  
QY 1636 TTTGGCTCTGTGAGCGAAGACGTAAAGAGGCACTGAGAGTCTGTGCGCTGCTCCGTT 1695  
Db 1741 TTTGGCTCTGTGAGCGAAGACGTAAAGAGGCACTGAGAGTCTGTGCGCTGCTCCGTT 1800  
QY 1696 CCCCTGGGGGTGAAGCAGAACTTCTGCACTGGGCTCTCTGTTGGGTCAAGACCTAAT 1755  
Db 1801 CCCCTGGGGGTGAAGCAGAACTTCTGCACTGGGCTCTCTGTTGGGTCAAGACCTAAT 1860  
QY 1756 GCGACCAACCCAGAGACACCTGGAAGCGCTTCCACTGTTTGGAGCTCAAGACAA 1815  
Db 1861 GCGACCAACCCAGAGACACCTGGAAGCGCTTCCACTGTTTGGAGCTCAAGACAA 1920  
QY 1816 GAGTTTGTGCTTGGCATTTAAACAGCTTCCAGAAATGTGGCTTCGATTTAACAGAAC 1875  
Db 1921 GAGTTTGTGCTTGGCATTTAAACAGCTTCCAGAAATGTGGCTTCGATTTAACAGAAC 1980  
QY 1876 CTGACCTTATGATGATCTTCTCTGCTCCAGACACTGTCCGATTTTGGGAAAAATTCG 1935  
Db 1981 CTGACCTTATGATGATCTTCTCTGCTCCAGACACTGTCCGATTTTGGGAAAAATTCG 2040  
QY 1936 GTGATGTCAAGAGGATCTTCCCAAGAGATGAGTCCGCTGAGGACATGTCCTGTGCTCCT 1995  
Db 2041 GTGATGTCAAGAGGATCTTCCCAAGAGATGAGTCCGCTGAGGACATGTCCTGTGCTCCT 2100  
QY 1996 CTATGATGCGGAGTAAAGACCTCATTTAGAGAGACGTGGGAAATTTCTGCTCATGCTT 2055  
Db 2101 CTATGATGCGGAGTAAAGACCTCATTTAGAGAGACGTGGGAAATTTCTGCTCATGCTT 2160  
QY 2056 GGCACCAACCAACACTGTCGAGCTGACCTGGGACAGACATCTGACAGAGCGGCGC 2115  
Db 2161 GGCACCAACCAACACTGTCGAGCTGACCTGGGACAGACATCTGACAGAGCGGCGC 2220  
QY 2116 ATGAAGACCTGTGTGCAAGCTGAGGACATCCCACTGCAAGAGATACAGACCTGTGATTT 2175  
Db 2221 ATGAAGACCTGTGTGCAAGCTGAGGACATCCCACTGCAAGAGATACAGACCTGTGATTT 2280

QY	2176	GGAAATGACAGATTAAACCCCTGGTGTGAGACACCTCTGAGAAATTCGTACATGCGCAACCGT	2235
Dp	2281	AGAAATGACAGATTAAACCCCTGGTGTGAGACACCTCTGAGAAATTCGTATGCGCAACCGT	2340
QY	2236	AACCTAAGATCCCTCAACCTTGGGAGGACCCACCTGAAGAAAGAGATGTAAAGATGCG	2295
Dp	2341	AACCTAAGATCCCTCAACCTTGGGAGGACCCACCTGAAGAAAGAGATGTAAAGATGCG	2400
QY	2296	TGTGAGCCTTAAACACCCAAATGTTTGTGAGTCTTTGAGGCTGGAATGCTGTGGA	2355
Dp	2401	TGTGAGCCTTAAACACCCAAATGTTTGTGAGTCTTTGAGGCTGGAATGCTGTGGA	2466
QY	2356	TTGACCCCATGCTGTACTGTAAAGATCTTCCCAATCCCTTAGACCTCCCCAGCCTGA	2415
Dp	2461	TTGACCCCATGCTGTACTGTAAAGATCTTCCCAATCCCTTAGACCTCCCCAGCCTGA	2520
QY	2416	TCTGTAGACCTGAGGAAACAGAGGTGACAGACCGAGGAGTAAATGCTCTCAGTGAATGC	2475
Dp	2521	TCTGTAGACCTGAGGAAACAGAGGTGACAGACCGAGGAGTAAATGCTCTCAGTGAATGC	2580
QY	2476	TTGAGAGTCTCCAGTGGCCCTGACAAAGCTGATATCGAGAGATGTGTGACATCAACGC	2535
Dp	2581	TTGAGAGTCTCCAGTGGCCCTGACAAAGCTGATATCGAGAGATGTGTGACATCAACGC	2640
QY	2536	ACGGGTGCCAGAGTCTTGCCCTCAGCCCTCGTACAGCAACCGGAGTTTGAACAACCTGTGC	2595
Dp	2641	ACGGGTGCCAGAGTCTTGCCCTCAGCCCTCGTACAGCAACCGGAGTTTGAACAACCTGTGC	2700
QY	2596	CTATCCAAACAAGCTGAGGGAACGAAGGTGTAAATCTACTGTGTGCATGATGAGGCTT	2655
Dp	2701	CTATCCAAACAAGCTGAGGGAACGAAGGTGTAAATCTACTGTGTGCATGATGAGGCTT	2760
QY	2656	CCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTCCACCTGGAACACGCTGAGCT	2715
Dp	2761	CCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTCCACCTGGAACACGCTGAGCT	2820
QY	2716	GGTTTCTTCTGCACTTGGCCCTTAATGGGTAACTCAATGGCTGAGGCACTGAGCCTTAGCATG	2775
Dp	2821	GGTTTCTTCTGCACTTGGCCCTTAATGGGTAACTCAATGGCTGAGGCACTGAGCCTTAGCATG	2880
QY	2776	AACCCGTGGAAGACAATGGGCGTGAAAGCTTCTGTGCGAGGATCATGAGAAACAATCTTGT	2835
Dp	2881	AACCCGTGGAAGACAATGGGCGTGAAAGCTTCTGTGCGAGGATCATGAGAAACAATCTTGT	2940
QY	2836	CATCTCCAGACCTGAGAGTGTAAAGTGCATCTCACCGCCGCGTGTGTGAGAGTCTG	2895
Dp	2941	CATCTCCAGACCTGAGAGTGTAAAGTGCATCTCACCGCCGCGTGTGTGAGAGTCTG	3000
QY	2896	TCTGTGTGATCTCGAGAGACACACTCTGAAGGCTGGAATCTTACCGAACATGCTCTG	2955
Dp	3001	TCTGTGTGATCTCGAGAGACACACTCTGAAGGCTGGAATCTTACCGAACATGCTCTG	3060
QY	2956	GGTACCGGTGGGGTTGCTGCGCTGTGCGAGGGAACGAAGAGTGTCTGACG	3015
Dp	3061	GGTACCGGTGGGGTTGCTGCGCTGTGCGAGGGAACGAAGAGTGTCTGACG	3120
QY	3016	AGACTCGGGTTGAAGGACATGTGGAACCTGATTTCTGATTTCTGTGAGGCACTCTCTTGGCC	3075
Dp	3121	AGACTCGGGTTGAAGGACATGTGGAACCTGATTTCTGATTTCTGTGAGGCACTCTCTTGGCC	3180
QY	3076	CTTTCTCTGCAACCGGCATCTGAACAGTCTAAACCTGTGCGAATTAATTGATGCCAA	3135
Dp	3181	CTTTCTCTGCAACCGGCATCTGAACAGTCTAAACCTGTGCGAATTAATTGATGCCAA	3240
QY	3136	GGAAATGAAGAGCTGTGTTTGCGGCTTTGCGCTGCGCAAGCTTAACCTTAACAATTAATGGG	3195
Dp	3241	GGAAATGAAGAGCTGTGTTTGCGGCTTTGCGCTGCGCAAGCTTAACCTTAACAATTAATGGG	3300
QY	3196	CTGTGGAATGGCACTACCTCTGTGCAATTAAGAAAGCTGTGAGGAAGTCACTACTC	3255
Dp	3301	CTGTGGAATGGCACTACCTCTGTGCAATTAAGAAAGCTGTGAGGAAGTCACTACTC	3360
QY	3256	AAGCCCGAGTGTAAATGACGATATTGGCAATCTTTTGAATGAAGATGACGCAAC	3312

DB	3361	AAAGCCGAGTCGTAAATTGACGGTAGTTGGCATTCTTTGATGAAAGTACCGGTAC	3417
RESULT 5			
AX478549			
LOCUS	AX478549	3489 bp	DNA
DEFINITION	Sequence 3 from Patent WO0248362.		linear
ACCESSION	AX478549		
VERSION	AX478549.1	GI:22217318	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
JOURNAL	1		
FEATURES	Ramkumar, J. and Arvizu, C.		
source	Embryogenesis associated proteins		
	Patent: WO 0248362-A 3 20-JUN-2002;		
	Incyte Genomics, Inc. (US)		
	Location/Qualifiers		
	1..3489		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 7474830CB1"		
ORIGIN			
Query Match	77.7%; Score 3133.6; DB 6; Length 3489;		
Best Local Similarity	95.2%; Pred. No. 0;		
Matches 3308; Conservative	0; Mismatches 4; Indels 162; Gaps 1;		
QY	1	ATGGAAGAGAGCAAAATGCTCAACCTTTTCCAGCTACGGGTGCAGATGTTCTCTAGAG	60
DB	1	ATGGAAGAGAGCAAAATGCTCAACCTTTTCCAGCTACGGGTGCAGATGTTCTCTAGAG	60
QY	61	CTAAGCAAGAGAAATTCAGACACTTAAGGAATTAATAAGAAATCTTCAGATCG	120
DB	61	CTAAGCAAGAGAAATTCAGACACTTAAGGAATTAATAAGAAATCTTCAGATCG	120
QY	121	ACGACATGCTCTATTCAGAGTTTGAATGAGAAATCCAAACGGAATGTCGGCACTC	180
DB	121	ACGACATGCTCTATTCAGAGTTTGAATGAGAAATCCAAACGGAATGTCGGCACTC	180
QY	181	CTCTTGATGATGATTAATGAGATCGCTGGCTGAGTACATTTAGCATTTTGA	240
DB	181	CTCTTGATGATGATTAATGAGATCGCTGGCTGAGTACATTTAGCATTTTGA	240
QY	241	AACATGAACCTGCGAACCTCTCTGAGAGAGCGACGGATGACATGAAAA	289
DB	241	AACATGAACCTGCGAACCTCTCTGAGAGAGCGACGGATGACATGAAAA	289
QY	290	-----	289
DB	301	GATCTGAAGCAACGATGACTGACCAAGACCAAGCAAGAAAAAGTCCAGAAATTA	360
QY	290	-----	289
DB	361	TATGCGATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAAC	420
QY	290	-----	318
DB	421	AAGTATTTGGAATTCATCTCTTTTTCAGAAATTTCAAGAGTATGAGAAAGAAAGT	480
QY	319	GGCAGACACAGAGAGAGAGAAAGAAATTTCAAGAGTATGAGAAAGAAAGT	378
DB	481	GGCAGACACAGAGAGAGAGAGAAAGAAATTTCAAGAGTATGAGAAAGAAAGT	540
QY	379	ACAGCAGACAGAGAGAGAGAAAGAAAGTATGAGAGTGAACATGAGGACTCAAGAGTCA	438
DB	541	ACAGCAGACAGAGAGAGAGAGAAAGAAAGTATGAGAGTGAACATGAGGACTCAAGAGTCA	600
QY	439	GTCATGACCAAAATTCGCTGAGAGAGAGATGATGTCGTAGTTTGAAGAACTGCTGCT	498

Db 601 GTGATGACCAATTCGTGAGAGAGGATGATGCTGTGATGTTGAAAACATCTCTCT 660  
Qy 499 GACTGGCCGGAAATCCAAACGTTGGCTGTGTCTTTTATTCAGACCGGTGGGCTTCCGG 558  
Db 661 GACTGGCCGGAAATCCAAACGTTGGCTGTGTCTTTTATTCAGACCGGTGGGCTTCCGG 720  
Qy 559 CCTGCGACGGTGTCTGTGACGGAAGTCAGGAATTTGGAAATCGGCTCTAGCCGAGAG 618  
Db 721 CCTGCGACGGTGTGTCTGTGACGGAAGTCAGGAATTTGGAAATCGGCTCTAGCCGAGAG 780  
Qy 619 ATCTGCTGTGTGGGCGCAAGTGGACTTTCACGGGAATGTTCTCTAGCTTCTTC 678  
Db 781 ATCTGCTGTGTGGGCGCAAGTGGACTTTCACGGGAATGTTCTCTAGCTTCTTC 840  
Qy 679 CTCCCGCTTAAAGATGCAACGCGGAAGAGAGAGCAATGTTCACGAATTCATCTCCAG 738  
Db 841 CTCCCGCTTAAAGATGCAACGCGGAAGAGAGAGCAATGTTCACGAATTCATCTCCAG 900  
Qy 739 GAGTGGCCGAGCTCCCGAGCTCCGGTGAACGAGATCATGTCCCGACCGAAAGGCTGTG 798  
Db 901 GAGTGGCCGAGCTCCCGAGCTCCGGTGAACGAGATCATGTCCCGACCGAAAGGCTGTG 960  
Qy 799 TTTCATATTGAAGGTTTGCATGACCTTGGGCTCTGTCTCAACAATGACACAAAGCTCTGC 858  
Db 961 TTTCATATTGAAGGTTTGCATGACCTTGGGCTCTGTCTCAACAATGACACAAAGCTCTGC 1020  
Qy 859 AAAAGCTGGGCTGAGAGACAGCTCTCGTTACCTTCATACGCAAGTCTGTGAGAGAGTTC 918  
Db 1021 AAAAGCTGGGCTGAGAGACAGCTCTCGTTACCTTCATACGCAAGTCTGTGAGAGAGTTC 1080  
Qy 919 CTGCTCCCTGAGTCTCTGTGATGTCACCGTCAAGAGAGTGGGCGACAGAGAGTCCAG 978  
Db 1081 CTGCTCCCTGAGTCTCTGTGATGTCACCGTCAAGAGAGTGGGCGACAGAGAGTCCAG 1140  
Qy 979 TCAAGAGTCTGTCTCCCGTTACCTGTATTAGAGAAATCTCCGGGAAACAAAGAAATC 1038  
Db 1141 TCAAGAGTCTGTCTCCCGTTACCTGTATTAGAGAAATCTCCGGGAAACAAAGAAATC 1200  
Qy 1039 CACTGTCTCTTGAAGCGCGGATTTGTGATGATCAATCAAGAGACAAAGGTTGCGTGCATC 1098  
Db 1201 CACTGTCTCTTGAAGCGCGGATTTGTGATGATCAATCAAGAGACAAAGGTTGCGTGCATC 1260  
Qy 1099 ATGAACAAACGCTGAGCTCTGACCAAGTCCAGAGGCGCGCGTGGGCTCTCTCATCTGC 1158  
Db 1261 ATGAACAAACGCTGAGCTCTGACCAAGTCCAGAGGCGCGCGTGGGCTCTCTCATCTGC 1320  
Qy 1159 GTGGCCCTGCAAGCTGCAAGAGAGTGTGGGAGAGAGCTGCGCCCTTTCACCAACGCTTC 1218  
Db 1321 GTGGCCCTGCAAGCTGCAAGAGAGTGTGGGAGAGAGCTGCGCCCTTTCACCAACGCTTC 1380  
Qy 1219 ACAGGCTGCAAGCGCGCTTTTGTGTTCATAGCTCACCCCTCGAGGCTGTGCGGCGC 1278  
Db 1381 ACAGGCTGCAAGCGCGCTTTTGTGTTCATAGCTCACCCCTCGAGGCTGTGCGGCGC 1440  
Qy 1279 TGTCTCATCTGAGGAAAGAGTTGTCTGAAGGCTCTGCGCGAATGGCTGTGAGGGA 1338  
Db 1441 TGTCTCATCTGAGGAAAGAGTTGTCTGAAGGCTCTGCGCGAATGGCTGTGAGGGA 1500  
Qy 1339 GTGTGAAATAGGAATCAGTGTGTGATGTGACGACCTCATAGTTCAGAGACTCGGAGAG 1398  
Db 1501 GTGTGAAATAGGAATCAGTGTGTGATGTGACGACCTCATAGTTCAGAGACTCGGAGAG 1560  
Qy 1399 TCTGAGCTCGTGTCTGTTCATAGGAATCTTCTTCCAGACAGCCACTGTGAGAG 1458  
Db 1561 TCTGAGCTCGTGTCTGTTCATAGGAATCTTCTTCCAGACAGCCACTGTGAGAG 1620  
Qy 1459 TACTACACCTTCTTCCAACTCAGTTCAGAGACTTCTGTGCGGCTTGTACTAGTGTGA 1518  
Db 1621 TACTACACCTTCTTCCAACTCAGTTCAGAGACTTCTGTGCGGCTTGTACTAGTGTGA 1680  
Qy 1519 GAGGCTGGAATCGAGCTGCTGCGCTGTGATGTTGAGAGACAAAGAGTCC 1578

Db 1681 GAGGCTGGAATCGAGCTGCTGCGCTGTGATGTTGAGAGACAAAGAGTCC 1740  
Qy 1579 ATGAGCTTAAACAGGAGCTTCCATATCCACTCGCTTTGAGAGCGTTTCTGTGTT 1638  
Db 1741 ATGAGCTTAAACAGGAGCTTCCATATCCACTCGCTTTGAGAGCGTTTCTGTGTT 1800  
Qy 1639 GGCTGTGAGCGAAGAGCTTAAGAGGCGCACTGAGGCTCTGTGAGGCTGTCCGTTCC 1698  
Db 1801 GGCTGTGAGCGAAGAGCTTAAGAGGCGCACTGAGGCTCTGTGAGGCTGTCCGTTCC 1860  
Qy 1699 CTGGGGGTGAAGCAAGACTTCTGCACTGGGTCTCTGTTGGGTGACGACCTTAATGCC 1758  
Db 1861 CTGGGGGTGAAGCAAGACTTCTGCACTGGGTCTCTGTTGGGTGACGACCTTAATGCC 1920  
Qy 1759 ACCACCCGAGAGCACCTGAGAGCGCTTCCACTGTCTTTTCAGACTCAAGACAAAGAG 1818  
Db 1921 ACCACCCGAGAGCACCTGAGAGCGCTTCCACTGTCTTTTCAGACTCAAGACAAAGAG 1980  
Qy 1819 TTTGTGCTTGGATTAACAGCTTCCAAAGATGTGGCTTCCGATTAACCAAGACCTG 1878  
Db 1981 TTTGTGCTTGGATTAACAGCTTCCAAAGATGTGGCTTCCGATTAACCAAGACCTG 2040  
Qy 1879 GACTTGAATGATCTTCTCTGTGCTCCAGACTGTCCGATTTGGGAAATTCGGGTG 1938  
Db 2041 GACTTGAATGATCTTCTCTGTGCTCCAGACTGTCCGATTTGGGAAATTCGGGTG 2100  
Qy 1939 GATGTCAAAGGAGATCTTCCAAAGATGAGTCCGTGAGGAGATCTCTGTGCTTCCCTA 1998  
Db 2101 GATGTCAAAGGAGATCTTCCAAAGATGAGTCCGTGAGGAGATCTCTGTGCTTCCCTA 2160  
Qy 1999 TGAATGCGGATTAAGACCTTCATTAAGAGAGAGTGGAAATTTCTGCTCATCTTGGC 2058  
Db 2161 TGAATGCGGATTAAGACCTTCATTAAGAGAGAGTGGAAATTTCTGCTCATCTTGGC 2220  
Qy 2059 ACCACCCGAGAGCACCTGAGAGCGCTTCCAGCTGAGAGATCTTCCAGAGCGGCGCATG 2118  
Db 2221 ACCACCCGAGAGCACCTGAGAGCGCTTCCAGCTGAGAGATCTTCCAGAGCGGCGCATG 2280  
Qy 2119 AAGACCTGTGTGCAAGCTGAGGATCCCACTGCGAAGATTAAGACCTGATGTTTGA 2178  
Db 2281 AAGACCTGTGTGCAAGCTGAGGATCCCACTGCGAAGATTAAGACCTGATGTTTGA 2340  
Qy 2179 AATGCAAGATTAACCTCTGTGTGCAAGCTTCTGAGAAATTCGATAGGCCAACCTTAAC 2238  
Db 2341 AATGCAAGATTAACCTCTGTGTGCAAGCTTCTGAGAAATTCGATAGGCCAACCTTAAC 2400  
Qy 2239 CTAAGATCTCTCAACTTGGAGAGAGCCACCTGAAGAGAGATGTAAGAGATGCGTGT 2298  
Db 2401 CTAAGATCTCTCAACTTGGAGAGAGCCACCTGAAGAGAGATGTAAGAGATGCGTGT 2460  
Qy 2299 GAAGCTTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGATTTGCTGTGATTTG 2358  
Db 2461 GAAGCTTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGATTTGCTGTGATTTG 2520  
Qy 2259 ACCCATGCTGTATCTGAGAGATCTCCCAATCTTTCAGACCTGCCAGGCTGAATCT 2418  
Db 2521 ACCCATGCTGTATCTGAGAGATCTCCCAATCTTTCAGACCTGCCAGGCTGAATCT 2580  
Qy 2419 CTGAGCTGTGAGAGAAACAGGTGACAGACCAAGGAGTAATGCTCTCATAGATGCTTG 2478  
Db 2581 CTGAGCTGTGAGAGAAACAGGTGACAGACCAAGGAGTAATGCTCTCATAGATGCTTG 2640  
Qy 2479 AGAGTCTCCAGTGCCTGTGCAAGAGCTGATTACTGAGAGACTGTGACATACAGCAAG 2538  
Db 2641 AGAGTCTCCAGTGCCTGTGCAAGAGCTGATTACTGAGAGACTGTGACATACAGCAAG 2700  
Qy 2539 GGTTCGAGAGTCTGAGCTGAGCTGAGCTGAGCAACGAGGCTTGAACAACCTGTGCTTA 2598  
Db 2701 GGTTCGAGAGTCTGAGCTGAGCTGAGCTGAGCAACGAGGCTTGAACAACCTGTGCTTA 2760  
Qy 2599 TCCAAACAAAGCTGTGGGAGAGAGGTGAATTACTGTGTGATCATAGAGCTTCCC 2658  
Db 2761 TCCAAACAAAGCTGTGGGAGAGAGGTGAATTACTGTGTGATCATAGAGCTTCCC 2820

OY	1	ATGGAAGGAGCAAAATGCGCTACACCTTTTCAGACTACGGGCTGCAAATGGTGTCTCATATG	60
DB	1	ATGGAAGGAGCAAAATGCGCTACCTTTTCAGACTACGGGCTGCAAATGGTGTCTCATATG	60
OY	61	CTAGAACAAGGAAGAAATTCACACATTCACAGGAATTCATAAAGAGAAATCTTCACAAATCG	120
DB	61	CTAGAACAAGGAAGAAATTCACACATTCACAGGAATTCATAAAGAGAAATCTTCACAAATCG	120
OY	121	ACCAATGCTCTAATTCACAGCTTTGAAATCGAAGATGCGAATGCGAAATGTCGCACTC	180
DB	121	ACCAATGCTCTAATTCACAGCTTTGAAATCGAAGATGCGAATGCGAAATGTCGCACTC	180
OY	181	CTCTTGAATGAGTATTAATGAGAGCATGCGTGGCTGAGGCTAGTCCATTAGCATCTTTGAA	240
DB	181	CTCTTGAATGAGTATTAATGAGAGCATGCGTGGCTGAGGCTAGTCCATTAGCATCTTTGAA	240
OY	241	AACATGAACCTGCGAACCCTCTCGGAGAAAGGCAACGGATGACATGAAAA-----	289
DB	241	AACATGAACCTGCGAACCCTCTCGGAGAAAGGCAACGGATGACATGAAAAATTCACAGAA	300
OY	290	-----	289
DB	301	GATCCTGAAGCAAGATGACTGACCAAGACCAAGCAAGAAAAAGTCAGAAAAATAA	360
OY	290	-----	289
DB	361	TATGGCATGACTAAGCTTAATCTTGGGGGTGTCTGACATCTGTGACTCGAATTAATAACAC	420
OY	290	-----AAATTCACAAGCTATGGAACAAAGGT	318
DB	421	AAGTATGTTGGAATTCATCTTCTTTTGCAAGAAATTCACAAGCTATGGAACAAAGGT	480
OY	319	GCCAAGAGAGAGAGACAGAAAGAACAGAAATTTCCACAAGCTATGGAACAAAGAGTGC	378
DB	481	GCCAAGAGAGAGAGACAGAAAGAACAGAAATTTCCACAAGCTATGGAACAAAGAGTGC	540
OY	379	ACAGACAGACAGACAGAAAGAACAGACATGAGAGTGCACATGGGACTACAAAGTCA	438
DB	541	ACAGACAGACAGACAGAAAGAACAGACATGAGAGTGCACATGGGACTACAAAGTCA	600
OY	439	GTCATGACCAAAATTCGTCGAGAGAGAGAGATGTAAGTGTAGTTTGAACAACCTGCTGT	498
DB	601	GTCATGACCAAAATTCGTCGAGAGAGAGAGATGTAAGTGTAGTTTGAACAACCTGCTGT	660
OY	499	GACTGACCGGAAATGCAAACTTTGGCTGTCTTTTGAATTCAGACCGTGGGGCTTCCGG	558
DB	661	GACTGACCGGAAATGCAAACTTTGGCTGTCTTTTGAATTCAGACCGTGGGGCTTCCGG	720
OY	559	CTGCGACCGGTGTTCTGACGGAAGTCAAGAAATGGGAAATCGGCTCTAGCCAGAAAG	618
DB	721	CCTGCGACCGGTGTTCTGACGGAAGTCAAGAAATGGGAAATCGGCTCTAGCCAGAAAG	780
OY	619	ATCGTGTCTGTCTGTGGGCGCAGAGTGTGACTCTACAGGGAATGTTCTCTAGCTTTCTTC	678
DB	781	ATCGTGTCTGTCTGTGGGCGCAGAGTGTGACTCTACAGGGAATGTTCTCTAGCTTTCTTC	840
OY	679	CTCCCGGTTAGAGAGATGACAGCGGAAGAGAGAGACAGTGTCAAGATTCATCTCCAGG	738
DB	841	CTCCCGGTTAGAGAGATGACAGCGGAAGAGAGAGACAGTGTCAAGATTCATCTCCAGG	900
OY	739	GAGTGGCCAGACTCCCAAGCTCCGGGTGACGGAATCATGTGCCGACCAAGAAAGCTGTGG	798
DB	901	GAGTGGCCAGACTCCCAAGCTCCGGGTGACGGAATCATGTGCCGACCAAGAAAGCTGTGG	960
OY	799	TTTCAATCATTTGACGGTTTGCATGACCTGGGCTGTGTCTCTCAACATGACACAAAGCTCTGC	858
DB	961	TTTCAATCATTTGACGGTTTGCATGACCTGGGCTGTGTCTCTCAACATGACACAAAGCTCTGC	1020
OY	859	AAAAGCTGGGCTGAGAGACAGCTCCGTTCAACCCGTAATACGACAGCTGCTGAGGAAGGTC	918
DB	1021	AAAAGCTGGGCTGAGAGACAGCTCCGTTCAACCCGTAATACGACAGCTGCTGAGGAAGGTC	1080
OY	919	CTGCTCCCTGAGTCTCTTCTGATGTCAACCGTCAAGAACGTGGGCAACAGAAAGCTCAAG	978



Db 1081 CTGCTCCCTGAGTCTTCTCTGATCGTCAACCGTCAGAGACGTGGGCACAGAGAGCTCAAG 1140  
Qy 979 TCAGAGGTGCGTGTCTCCCGTTACTGTATTAGTAAGAAATCTCCGGGGAACAAAGATTC 1038  
Db 1141 TCAGAGGTGCGTGTCTCCCGTTACTGTATTAGTAAGAAATCTCCGGGGAACAAAGATTC 1200  
Qy 1039 CACTTGTCTTGTAGACGCGGAGATTGGTAGCATCAGAAAGACAAGAGGTGGGTGTCATC 1098  
Db 1201 CACTTGTCTTGTAGACGCGGAGATTGGTAGCATCAGAAAGACAAGAGGTGGGTGTCATC 1260  
Qy 1099 ATGAACAACCGTAGCTGTCTGACCAAGTGCAGAGTGCCTGCGGTGAGTCTTCTCATCTGC 1158  
Db 1261 ATGAACAACCGTAGCTGTCTGACCAAGTGCAGAGTGCCTGCGGTGAGTCTTCTCATCTGC 1320  
Qy 1159 GTGGCCCTGACAGTGCAGAGACGTGGTGGGGAGACCGTGCCTCTTCAACCAACGCTC 1218  
Db 1321 GTGGCCCTGACAGTGCAGAGACGTGGTGGGGAGACCGTGCCTCTTCAACCAACGCTC 1380  
Qy 1219 ACAGGCTGTGACGCGCTTTTGTGTTCATCAGCTCACCCCTGAGGCGTGGTCCGCGC 1278  
Db 1381 ACAGGCTGTGACGCGCTTTTGTGTTCATCAGCTCACCCCTGAGGCGTGGTCCGCGC 1440  
Qy 1279 TGTCTCAATCTGAGAGAAAGATTGTCTTGAAGCGCTTCTGCGTATGCTGTGAGGGA 1338  
Db 1441 TGTCTCAATCTGAGAGAAAGATTGTCTTGAAGCGCTTCTGCGTATGCTGTGAGGGA 1500  
Qy 1339 GTGTGGAATPAGGAATCAGTGTGTAAGTGAACGACCTCATGTTCAGAGATCGGGGAG 1398  
Db 1501 GTGTGGAATPAGGAATCAGTGTGTAAGTGAACGACCTCATGTTCAGAGATCGGGGAG 1560  
Qy 1399 TCTGAGCTCCGTGTCTGTTCACATGAAATCTTCTCCAGACAGCCACTGTGAGAG 1458  
Db 1561 TCTGAGCTCCGTGTCTGTTCACATGAAATCTTCTCCAGACAGCCACTGTGAGAG 1620  
Qy 1459 TACTACACCTTCTTCACTCAGTCTCAAGACCTTGTGTGCGCTTGTACTAGTGTTA 1518  
Db 1621 TACTACACCTTCTTCACTCAGTCTCAAGACCTTGTGTGCGCTTGTACTAGTGTTA 1680  
Qy 1519 GAGGGCCGTGGAATGAGACCAAGCTCTGCGCTGTGTAAGTGAAGAAACAAAGAGTCC 1578  
Db 1681 GAGGGCCGTGGAATGAGACCAAGCTCTGCGCTGTGTAAGTGAAGAAACAAAGAGTCC 1740  
Qy 1579 ATGAGCTTTAAACAGGACAGCTTCCATATCCACTGCTTGGATGAAAGCTTTCTTGT 1638  
Db 1741 ATGAGCTTTAAACAGGACAGCTTCCATATCCACTGCTTGGATGAAAGCTTTCTTGT 1800  
Qy 1639 GAGCTCTGAGGAAAGAGTAAAGAGGCCACTGAGAGTCTGTGGGCTGTCCGTTCC 1698  
Db 1801 GAGCTCTGAGGAAAGAGTAAAGAGGCCACTGAGAGTCTGTGGGCTGTCCGTTCC 1860  
Qy 1699 CTGGGGGGTGAAGCAAGCTTCTGACAGTGGGTCTCTGTGGGTCAACAGCTTAATGCC 1758  
Db 1861 CTGGGGGGTGAAGCAAGCTTCTGACAGTGGGTCTCTGTGGGTCAACAGCTTAATGCC 1920  
Qy 1759 ACCACCCAGAGACACCTGTGACGCTTCCACTGTCTTTCAGAGACTCAAGACAAAGAG 1818  
Db 1921 ACCACCCAGAGACACCTGTGACGCTTCCACTGTCTTTCAGAGACTCAAGACAAAGAG 1980  
Qy 1819 TTTGTTCGCTTGGCATTTAAACAGCTTCCAGAGTGTGGCTTCCGATTAACAGAACTG 1878  
Db 1981 TTTGTTCGCTTGGCATTTAAACAGCTTCCAGAGTGTGGCTTCCGATTAACAGAACTG 2040  
Qy 1879 GACTTGATAGACATCTTCTTCTGCTCTGACAGCTGTCCGTAATTTGCGGAAATTCGGGTG 1938  
Db 2041 GACTTGATAGACATCTTCTTCTGCTCTGACAGCTGTCCGTAATTTGCGGAAATTCGGGTG 2100  
Qy 1939 GATGTCAAGAGGATTTTCCCAAGAGATGTCCGCTGAGCATGTCTGTGGTCCCTCTA 1998  
Db 2101 GATGTCAAGAGGATTTTCCCAAGAGATGTCCGCTGAGCATGTCTGTGGTCCCTCTA 2160  
Qy 1999 TGGATGCGGAGTAAAGACCTTCAATTGAGAGACGTGGAAATTTCTGCTCAAGTGTGGC 2058  
|||||

Db 2161 TGGATGCGGAGTAAAGACCTTCAATTGAGAGACGTGGAAAGATTTCTGCTCAAGTGTGGC 2220  
Qy 2059 ACCACCCAGACCTGCGGACAGTGTGACCTGGGACAGACATCTGACAGAGCGGGCCATG 2118  
Db 2221 ACCACCCAGACCTGCGGACAGTGTGACCTGGGACAGACATCTGACAGAGCGGGCCATG 2280  
Qy 2119 AAGACCTGTGTGCAAGCTGAGGACATCCCACTGCAAGATACAGACCTGTATGTTAGA 2178  
Db 2281 AAGACCTGTGTGCAAGCTGAGGACATCCCACTGCAAGATACAGACCTGTATGTTAGA 2340  
Qy 2179 AATGACAGATTTACCCCTGTGTGACAGCTCTGAGAAATCGTCAATGSCCAACCTTAAC 2238  
Db 2341 AATGACAGATTTACCCCTGTGTGACAGCTCTGAGAAATCGTCAATGSCCAACCTTAAC 2400  
Qy 2239 CTAAGATCCCTCACTTGGGAGGACCACTGAAAGAAAGATGTAAGATGGCGTGT 2298  
Db 2401 CTAAGATCCCTCACTTGGGAGGACCACTGAAAGAAAGATGTAAGATGGCGTGT 2460  
Qy 2299 GAAACCTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGAAATGCTGTGAAATG 2358  
Db 2461 GAAACCTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGAAATGCTGTGAAATG 2520  
Qy 2359 ACCATGCTGTATCTGAGAGATCTCCCAATCTTACGACCTTCCCAAGCTGAAATCT 2418  
Db 2521 ACCATGCTGTATCTGAGAGATCTCCCAATCTTACGACCTTCCCAAGCTGAAATCT 2580  
Qy 2419 CTGAGCTGTGAGAGAAACAAGGTGACAGACAGGAGATTAATGCTCTCAATGATGCTTG 2478  
Db 2581 CTGAGCTGTGAGAGAAACAAGGTGACAGACAGGAGATTAATGCTCTCAATGATGCTTG 2640  
Qy 2479 AAGATCTCCAGTGGCGCTGTGACAGACCTGATCTGAGAGACTGTGACATCACAGCACG 2538  
Db 2641 AAGATCTCCAGTGGCGCTGTGACAGACCTGATCTGAGAGACTGTGACATCACAGCACG 2700  
Qy 2539 GGTGTCCAGAGATCTGGCTTCAAGCTCTGTGACGAAACCGAGCTTGAACACACTGTGCTTA 2598  
Db 2701 GGTGTCCAGAGATCTGGCTTCAAGCTCTGTGACGAAACCGAGCTTGAACACACTGTGCTTA 2760  
Qy 2599 TCCAAACAACGCTGGGGAAACGAAGTGTAAATCTACGTGTGATCCATGAGGCTTCCC 2658  
Db 2761 TCCAAACAACGCTGGGGAAACGAAGTGTAAATCTACGTGTGATCCATGAGGCTTCCC 2820  
Qy 2659 CACTGTAGTGTGACAGGCTGATGCTGAATCAGTGCACCTGACACAGGCTGTGTGT 2718  
Db 2821 CACTGTAGTGTGACAGGCTGATGCTGAATCAGTGCACCTGACACAGGCTGTGTGT 2880  
Qy 2719 TTTCTTGAACCTTGGCTTAATGGTAACTCATGAGTGAACGCACTGAGCTTTAGATGAAAC 2778  
Db 2881 TTTCTTGAACCTTGGCTTAATGGTAACTCATGAGTGAACGCACTGAGCTTTAGATGAAAC 2940  
Qy 2779 CCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACATCTGTCAAT 2838  
Db 2941 CCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTCATGAGAGAACATCTGTCAAT 3000  
Qy 2839 CTCAGAGACCTGAGATTGTAAAGTGCATCTCACCGCGGTGTGTGAGAGTCTGTCC 2898  
Db 3001 CTCAGAGACCTGAGATTGTAAAGTGCATCTCACCGCGGTGTGTGAGAGTCTGTCC 3060  
Qy 2899 TGTGTGATCTGAGAGACAGACACTGAAGAGCTGTGATCTCAGGACATATGCTCTGGGT 2958  
Db 3061 TGTGTGATCTGAGAGACAGACACTGAAGAGCTGTGATCTCAGGACATATGCTCTGGGT 3120  
Qy 2959 GACGATGGGGTGTGTGCGTGTGAGAGGACTGAAGAGAAAGACAGTGTCTGACAGAGA 3018  
Db 3121 GACGATGGGGTGTGTGCGTGTGAGAGGACTGAAGAGAAAGACAGTGTCTGACAGAGA 3180  
Qy 3019 CTGGGTTGAAAGGATGTGACATCTTGAATTCGTGTGAGGACATCTCTTGGCCCTT 3078  
Db 3181 CTGGGTTGAAAGGATGTGACATCTTGAATTCGTGTGAGGACATCTCTTGGCCCTT 3240  
Qy 3079 TCTGTCAAACCGGACATCTGACCAAGTCTAAACCTGTGTGAGAAATTAATTCAATGCCAAAGGA 3138  
Db 3241 TCTGTCAAACCGGACATCTGACCAAGTCTAAACCTGTGTGAGAAATTAATTCAATGCCAAAGGA 3300  
|||||





1438 CCAAGACCACTGTGAGAGTACTACACTTCTTCCACTGATCTCCAGAGACTTCTGT 1497  
3535 CCAAGACCACTGTGAGAGTACTACACTTCTTCCACTGATCTCCAGAGACTTCTGT 3594  
1498 GCCGCTTGTACTACGTGTGAGAGGCTCGAATAATGAGCCACTCTCTCTCTGTAC 1557  
3595 GCCGCTTGTACTACGTGTGAGAGGCTCGAATAATGAGCCACTCTCTCTCTGTAC 3654  
1558 GTTGAAGAGCAAAAGGTCATGAGGCTTAAAGAGGAGGCTTCATATCCACTGCTT 1617  
3655 GTTGAAGAGCAAAAGGTCATGAGGCTTAAAGAGGAGGCTTCATATCCACTGCTT 3714  
1618 TGGATGAAGCGTTTCTTGTGTCCTCGTGAAGCAAGCTTAAGAGGCTCCTGAGGTC 1677  
3715 TGGATGAAGCGTTTCTTGTGTCCTCGTGAAGCAAGCTTAAGAGGCTCCTGAGGTC 3774  
1678 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTCTCTG 1737  
3775 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTCTCTG 3834  
1738 TTGGGTCAAGCGCTTAATGCCACCAAGAGAGCAAGCTTGAAGGCTTCCACTGCTT 1797  
3835 TTGGGTCAAGCGCTTAATGCCACCAAGAGAGCAAGCTTGAAGGCTTCCACTGCTT 3894  
1798 TTGAGACTCAAGACAAAGAGTTTGTGCTTGGCAATTAAACAGCTTCCAAAGAGTGTG 1857  
3895 TTGAGACTCAAGACAAAGAGTTTGTGCTTGGCAATTAAACAGCTTCCAAAGAGTGTG 3954  
1858 CTTCGAGTTAACAGAACCTGAGCTTGATAGCATCTTCTCTGCTCCAGACTGTCCG 1917  
3955 CTTCGAGTTAACAGAACCTGAGCTTGATAGCATCTTCTCTGCTCCAGACTGTCCG 4014  
1918 TATTTGCGGAAATTCGGGTGATGTCAAGGATCTTCCAAAGAGTGAAGCGCTAG 1977  
4015 TATTTGCGGAAATTCGGGTGATGTCAAGGATCTTCCAAAGAGTGAAGCGCTAG 4074  
1978 GCATGCTCTGTGTCCTCTATGATGCGGAGTAAGACCTTCAATGAGAGAGTGGAA 2037  
4075 GCATGCTCTGTGTCCTCTATGATGCGGAGTAAGACCTTCAATGAGAGAGTGGAA 4134  
2038 GATTTCTGCTCCATGCTTGGCAACCAACCACTGCGGAGCTGAGCCTTGGGAGCAGC 2097  
4135 GATTTCTGCTCCATGCTTGGCAACCAACCACTGCGGAGCTGAGCCTTGGGAGCAGC 4194  
2098 ATCCGAGAGAGCGGGCCATGAGAGCCTGTGTGCAAGCTGAGGCAATCCCACTGCAAG 2157  
4195 ATCCGAGAGAGCGGGCCATGAGAGCCTGTGTGCAAGCTGAGGCAATCCCACTGCAAG 4254  
2158 ATACAGACCCCTGATGTTAGAAATGACAGATTAACCTTGTGTGTCAGACCTTGGAGA 2217  
4255 ATACAGACCCCTGATGTTAGAAATGACAGATTAACCTTGTGTGTCAGACCTTGGAGA 4314  
2218 ATGCTGATGCGCAACCGTAACTTAAGATCCCTCAACTTGGAGGAGCAACCTGAAGAA 2277  
4315 ATGCTGATGCGCAACCGTAACTTAAGATCCCTCAACTTGGAGGAGCAACCTGAAGAA 4374  
2278 GAGGATGTAAGATGCGGTGTAAGCTTAAACCAACCAATGTTTGTGTAAGTCTTGG 2337  
4375 GAGGATGTAAGATGCGGTGTAAGCTTAAACCAACCAATGTTTGTGTAAGTCTTGG 4434  
2338 AGGCTGATGCTGTGATGTAAGCCATGCTGTTAACCTGAAGATCTCCCAATCTTACG 2397  
4435 AGGCTGATGCTGTGATGTAAGCCATGCTGTTAACCTGAAGATCTCCCAATCTTACG 4494  
2398 ACCTCCCGCAGCGTAAATCTCTGAGCTTGGAGGAAACAAAGTGAAGAGCCAGGAGTA 2457  
4495 ACCTCCCGCAGCGTAAATCTCTGAGCTTGGAGGAAACAAAGTGAAGAGCCAGGAGTA 4554  
2458 ATGCTCTCAAGTGAAGCTTGAAGTCTTCCAGTGGCGCTTGCAGAACTGATATCTGAG 2517  
4555 ATGCTCTCAAGTGAAGCTTGAAGTCTTCCAGTGGCGCTTGCAGAACTGATATCTGAG 4614

2518 GACTGTGGCATCAGAGCCAGGGTGGCAGAGTCTGGCTTCAGGCTTCGTGAGAACCGG 2577  
4615 GACTGTGGCATCAGAGCCAGGGTGGCAGAGTCTGGCTTCAGGCTTCGTGAGAACCGG 4674  
2578 AGCTTGAACACCTGTGCTTATCCAAACAGAGCTTGGGAAACGAAGTGAATCTACTG 2637  
4675 AGCTTGAACACCTGTGCTTATCCAAACAGAGCTTGGGAAACGAAGTGAATCTACTG 4734  
2638 TGTGATCATGAGAGCTTCCCACTGATGTGAGAGAGCTGATGCTGAATCAGTGCAC 2697  
4735 TGTGATCATGAGAGCTTCCCACTGATGTGAGAGAGCTGATGCTGAATCAGTGCAC 4794  
2698 CTGACACGCGCTGCTGTGTTTCTTGAACCTTGCCTTAATGAGTAACTCATGCTGACG 2757  
4795 CTGACACGCGCTGCTGTGTTTCTTGAACCTTGCCTTAATGAGTAACTCATGCTGACG 4854  
2758 CACTGAGCTTGAACAGACCTTGTGGAAGCAATGAGCGTGAAGCTTCTGTGGAAGTC 2817  
4855 CACTGAGCTTGAACAGACCTTGTGGAAGCAATGAGCGTGAAGCTTCTGTGGAAGTC 4914  
2818 ATGAGAGAACCATCTTGCATCTCAGAGACCTGAGATTGTAAGTGCATCTCACGCC 2877  
4915 ATGAGAGAACCATCTTGCATCTCAGAGACCTGAGATTGTAAGTGCATCTCACGCC 4974  
2878 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGAGACAGACCTGAAGAGCTGAT 2937  
4975 GCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGAGACAGACCTGAAGAGCTGAT 5034  
2938 CTACAGAGCAATGCCCTTGGGTGAGAGGTTGCTGCGCTGTGCGAGGACTGAAGCA 2997  
5035 CTACAGAGCAATGCCCTTGGGTGAGAGGTTGCTGCGCTGTGCGAGGACTGAAGCA 5094  
2998 AAGAAGAGTGTCTGAGAGACTCGGGTGAAGGATGTGAGCTGACTTCTGATTTGCTGT 3057  
5095 AAGAAGAGTGTCTGAGAGACTCGGGTGAAGGATGTGAGCTGACTTCTGATTTGCTGT 5154  
3058 GAGGACCTTCTTGGCTTCTTCTGCAACCGGATCTGACAGTCTAAACCTGTGTGAG 3117  
5155 GAGGACCTTCTTGGCTTCTTCTGCAACCGGATCTGACAGTCTAAACCTGTGTGAG 5214  
3118 AATACTTCACTGCCAAGAGATGATGAAGTGTGTGCTTGGCTTGTGCTGCTCCAGTCT 3177  
5215 AATACTTCACTGCCAAGAGATGATGAAGTGTGTGCTTGGCTTGTGCTGCTCCAGTCT 5274  
3178 AACTTACAGATTAATGG 3194  
5275 AACTTACAGATTAATGG 5291

RESULT 8  
AX459873  
LOCUS AX459873 6939 bp DNA linear PAT 30-NOV-2002  
DEFINITION Sequence 16 from Patent WO0240666.  
ACCESSION AX459873  
VERSION AX459873.1 GI:21725645  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Techop, J. and Martignon, F.  
Proteins and dna sequences underlying these proteins used for  
treating inflammations  
Patent: WO 0240668-A 16 23-MAY-2002;  
Apotech Research and Development Ltd. (CH)  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..6939  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="NALP5/Fy8.dna /GENSCAN\_predicted\_CDS\_1/ 6939\_bp"

ORIGIN

Query Match 75.8%; Score 3057.8; DB 6; Length 6939;  
Beet Local Similarity 96.2%; Pred. No. 0;  
Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;

1 ATGGAAGAGCAAAATCGCTCACTTTTCAGCTACGGGCTGCAATGTGTCTATGAG 60  
3055 ATGGAAGAGCAAAATCGCTCACTTTTCAGCTACGGGCTGCAATGTGTCTATGAG 3114  
61 CTAGACAAAGGAATTTCAAGATTCAAGAAATTCTAAGAAATCTTCAAGATCG 120  
3115 CTAGACAAAGGAATTTCAAGATTCAAGAAATTCTAAGAAATCTTCAAGATCG 3174  
121 ACCAATGCTCTATTCCACAGTTTAAATCGAAGATGCCAAGTGAATGTCTGACATC 180  
3175 ACCAATGCTCTATTCCACAGTTTAAATCGAAGATGCCAAGTGAATGTCTGACATC 3234  
181 CTCTTGATGATGATTTATGAGCATCGCTGGCTGCTACCTCATTTAGATCTTTGAA 240  
3235 CTCTTGATGATGATTTATGAGCATCGCTGGCTGCTACCTCATTTAGATCTTTGAA 3294  
241 AACATGAACCTGCGAACTCTCTGCGAAGAGCAAGGATGACATGAAA----- 289  
3295 AACATGAACCTGCGAACTCTCTGCGAAGAGCAAGGATGACATGAAAAGATTCACCA 3354  
290 ----- 289  
3355 GAAGATCTGAAAGCAAGATGATGACAGCAAGCAAGCAAGAAAGTGCAGGAATT 3414  
290 -----AAATTTCA 297  
3415 TCACAGACTGTGCAACAAATAGTGCCACAGCTGCAGAGACAAAGAAACAAATTTCA 3474  
298 CAAGTATGGAACAAAGAGTGCACACAGACAGAGACAAAGAAATTTTCAAA 357  
3475 CAAGTATGGAACAAAGAGTGCACACAGACAGAGACAAAGAAATTTTCAAA 3534  
358 GCTATGGAACAAAGAGTGCACACAGACAGAGACAAAGAAATTTGAGATGAC 417  
3535 GCTATGGAACAAAGAGTGCACACAGACAGAGACAAAGAAATTTGAGATGAC 3594  
418 ACATGAGACTACAAAGATGACATGATGACCAATTCGCTGAGAGAGATGATAGTGT 477  
3595 ACATGAGACTACAAAGATGACATGATGACCAATTCGCTGAGAGAGATGATAGTGT 3654  
478 AGTTTGAACCACTGCTGCTGACCTGCGAAGATGCAACGTTGCTGTCTTTGAT 537  
3655 AGTTTGAACCACTGCTGCTGACCTGCGAAGATGCAACGTTGCTGTCTTTGAT 3714  
538 TCAGACCGGTGGGGCTTCCGGCTTCGACAGGTGTTCTGACAGAAATTTGGG 597  
3715 TCAGACCGGTGGGGCTTCCGGCTTCGACAGGTGTTCTGACAGAAATTTGGG 3774  
598 AAATGAGCTCTAGCAGAAAGATGCTGCTGCTGGGCGCAAGGTGATCTTACCAAGGA 657  
3775 AAATGAGCTCTAGCAGAAAGATGCTGCTGCTGGGCGCAAGGTGATCTTACCAAGGA 3834  
658 ATGTTCTCTAGCTTTCTTCTCCCGTTAGAGATGACAGCGAAGAGAGACAGT 717  
3835 ATGTTCTCTAGCTTTCTTCTCCCGTTAGAGATGACAGCGAAGAGAGACAGT 3894  
718 GTCAAGAGTTTATCTTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGAGATCATG 777  
3895 GTCAAGAGTTTATCTTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGAGATCATG 3954  
778 TCCCGACCAAGAGGCTGTTGTTTCAATGACGGTTTCATGACCTGGGCTCTGCTGC 837  
3955 TCCCGACCAAGAGGCTGTTGTTTCAATGACGGTTTCATGACCTGGGCTCTGCTGC 4014  
838 AACATGACAAAGCTCTGCAAGACCTGGGCTGAGAGACAGCTCCGTTCACTCTATA 897  
4015 AACATGACAAAGCTCTGCAAGACCTGGGCTGAGAGACAGCTCCGTTCACTCTATA 4074

QY 898 CGAGCTGTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACCGTCAGAGAC 957  
DB 4075 CGAGCTGTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACCGTCAGAGAC 4134  
QY 998 GTGGGCAAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTTACCTGTATGAGAGA 1017  
DB 4135 GTGGGCAAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTTACCTGTATGAGAGA 4194  
QY 1018 ATCTCCGGGAAACAAAGATTCATCTTCTCTTGAGCCGGAGATTGTGAGATCAAG 1077  
DB 4195 ATCTCCGGGAAACAAAGATTCATCTTCTCTTGAGCCGGAGATTGTGAGATCAAG 4254  
QY 1078 ACACAGGTTGCGGATCATGAAACACGTCGCTGTGACACAGTCCAGGTGCCC 1137  
DB 4255 ACACAGGTTGCGGATCATGAAACACGTCGCTGTGACACAGTCCAGGTGCCC 4314  
QY 1138 GCCGTGGCTCTCTCATCTGCTGCTGCTGACAGTGCAGAGACGTGTGGGAGAGCGTC 1197  
DB 4315 GCCGTGGCTCTCTCATCTGCTGCTGCTGACAGTGCAGAGACGTGTGGGAGAGCGTC 4374  
QY 1198 GCCCTTCAACCAAGCTCAAGGCTGCAAGCGCTTTTGTGTTTCACTAGCTCAC 1257  
DB 4375 GCCCTTCAACCAAGCTCAAGGCTGCAAGCGCTTTTGTGTTTCACTAGCTCAC 4434  
QY 1258 CCTGAGGCGTGTCCGGGCTGCTCAATGAGAGAAAGTGTCTGAAAGGCTTC 1317  
DB 4435 CCTGAGGCGTGTCCGGGCTGCTCAATGAGAGAAAGTGTCTGAAAGGCTTC 4494  
QY 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGAGTGCAGCTC 1377  
DB 4495 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGAGTGCAGCTC 4554  
QY 4555 ATGTTCAAGACCTCGGAGAGTGTGAGCTCGTGTCTGTTTCACTAATAAATCTTCTC 4614  
DB 1438 CCAGACCCCTGAGAGATGATCAACCTTTTCCACTCAAGTCTCAGAGCTTCTGT 1497  
4615 CCAGACCCCTGAGAGATGATCAACCTTTTCCACTCAAGTCTCAGAGCTTCTGT 4674  
QY 1498 GCCGCTTGTACTAGTGTGAGAGGCTGTGAAATCGAGCAGCTCTGCGCTGTGAC 1557  
DB 4675 GCCGCTTGTACTAGTGTGAGAGGCTGTGAAATCGAGCAGCTCTGCGCTGTGAC 4734  
QY 1558 GTTGAAGACAAAGAGTTCATGAGAGTTCATCAAGGAGCTTCAATATCACTGCTT 1617  
DB 4735 GTTGAAGACAAAGAGTTCATGAGAGTTCATCAAGGAGCTTCAATATCACTGCTT 4794  
QY 1618 TGGATGAAGCGTTTCTTGTGCTGCTGAGCCGAAGCTTAAGAGGCCATGAGAGTTC 1677  
DB 4795 TGGATGAAGCGTTTCTTGTGCTGCTGAGCCGAAGCTTAAGAGGCCATGAGAGTTC 4854  
QY 1678 CTGCTGGGCTGTCCCGTTCCCTGGGAGGTGAAGAGAGTTCGCACTGGGCTCTCTG 1737  
DB 4855 CTGCTGGGCTGTCCCGTTCCCTGGGAGGTGAAGAGAGTTCGCACTGGGCTCTCTG 4914  
QY 1738 TTGGGTGAGAGCTTAATGACCAACCCAGAGAGCACTTGAAGCGCTTCACTGTCTT 1797  
DB 4915 TTGGGTGAGAGCTTAATGACCAACCCAGAGAGCACTTGAAGCGCTTCACTGTCTT 4974  
QY 1798 TTGAGAGCTCAAGACAAAGATTTGTGCTTGGCATTTAAACAGTTTCAAGAGTGG 1857  
DB 4975 TTGAGAGCTCAAGACAAAGATTTGTGCTTGGCATTTAAACAGTTTCAAGAGTGG 5034  
QY 1858 CTTCGATTAACCAAGACCTGAGCTTGAATAGATCTTCTGCTGCTCAGACCTGCTCG 1917  
DB 5035 CTTCGATTAACCAAGACCTGAGCTTGAATAGATCTTCTGCTGCTCAGACCTGCTCG 5094  
QY 1918 TATTTGCGAAATTTGGGTGATGTCAAGAGGATTTTCCAGAGATGATCGCTGAG 1977  
DB 5095 TATTTGCGAAATTTGGGTGATGTCAAGAGGATTTTCCAGAGATGATCGCTGAG 5154  
QY 1978 GATGTCTGTGTCTCTATGATGAGTGGGATTAAGACCTTCATTTGAGAGACAGTGGAA 2037

```
Db 5155 GCATGCTCTGTCCTCTCTATGATGCGGGATAAGACCTCATATGAGAGACAGTGGAA 5214
Qy 2038 GATTTCCTGCTCAGTCTTGAGACCCACACACCGCGGAGCGTGGACCTGGGACGACG 2097
Db 5215 GATTTCTGCTCCAGTCTTGAGACCCACACACCGCGGAGCGTGGACCTGGGACGACG 5274
Qy 2098 ATCTGAGAGAGCGGGCATGAGACCTGTGTGCGCAAGTGGAGCATCCACCTGCAAG 2157
Db 5275 ATCTGAGAGAGCGGGCATGAGACCTGTGTGCGCAAGTGGAGCATCCACCTGCAAG 5334
Qy 2158 ATACAGACCTGATGTTAAGAAATGACAGATTACCTGCTGTGAGACCTCTGAGA 2217
Db 5335 ATACAGACCTGATGTTAAGAAATGACAGATTACCTGCTGTGAGACCTCTGAGA 5394
Qy 2218 ATTCGATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGGACCACTGAGAA 2277
Db 5395 ATTCGATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGGACCACTGAGAA 5454
Qy 2278 GAGGATGTAAGGATGCGTGTGAGACCTTAAACACCCAAATGTTTGTGAGTCTTGG 2337
Db 5455 GAGGATGTAAGGATGCGTGTGAGACCTTAAACACCCAAATGTTTGTGAGTCTTGG 5514
Qy 2338 AGGCTGATGCTGTGATGATGACCATGCTGTTACTGAAAGATCTCCAAATCTTAAG 2397
Db 5515 AGGCTGATGCTGTGATGATGACCATGCTGTTACTGAAAGATCTCCAAATCTTAAG 5574
Qy 2398 ACCCTCCCGAGCGTGAATCTCTGAGCCGTGGAGGAAACAAAGTGAAGACAGACAGGAA 2457
Db 5575 ACCCTCCCGAGCGTGAATCTCTGAGCCGTGGAGGAAACAAAGTGAAGACAGACAGGAA 5634
Qy 2458 ATGCTCTCAGTGTATGCTGAGAGTCTCCAGTGCAGCCCTGACAGAGCTGATCTGAG 2517
Db 5635 ATGCTCTCAGTGTATGCTGAGAGTCTCCAGTGCAGCCCTGACAGAGCTGATCTGAG 5694
Qy 2518 GACTGTGCAATCAAGCCACGGGTTGCCAGAGTCTGCGCTGACGCCCTGTCAGACCGG 2577
Db 5695 GACTGTGCAATCAAGCCACGGGTTGCCAGAGTCTGCGCTGACGCCCTGTCAGACCGG 5754
Qy 2578 AGCTTGAACAACCTGTGCTATCCAAACAAAGCCGGGGAACGAAGGTGAATCTACTG 2637
Db 5755 AGCTTGAACAACCTGTGCTATCCAAACAAAGCCGGGGAACGAAGGTGAATCTACTG 5814
Qy 2638 TGTGATCATGAGGCTTCCCACTGTAGTGTGAGAGGCTGATCTGATCAAGTGCAC 2697
Db 5815 TGTGATCATGAGGCTTCCCACTGTAGTGTGAGAGGCTGATCTGATCAAGTGCAC 5874
Qy 2698 CTGGAACGCGCTGCTGCTGTTTCTTGCACTTGCCTTAAGGTAATCTCATGCTGACG 2757
Db 5875 CTGGAACGCGCTGCTGCTGTTTCTTGCACTTGCCTTAAGGTAATCTCATGCTGACG 5934
Qy 2758 CACCTGAGCCTTGAAGAACCCGTGTGAAGCAATGGCGTGAAGCTCTGTGGGAGGTC 2817
Db 5935 CACCTGAGCCTTGAAGAACCCGTGTGAAGCAATGGCGTGAAGCTCTGTGGGAGGTC 5994
Qy 2818 ATGAGAGAACCATCTTGTCACTCCAGACCTGAGAGTGTAAAGTGTCACTCAACGCC 2877
Db 5995 ATGAGAGAACCATCTTGTCACTCCAGACCTGAGAGTGTAAAGTGTCACTCAACGCC 6054
Qy 2878 GCGTGTGTGAGAGTCTGCTGCTGTGATCTGAGAGAGCAACCTGAAGAGCTTGAT 2937
Db 6055 GCGTGTGTGAGAGTCTGCTGCTGTGATCTGAGAGAGCAACCTGAAGAGCTTGAT 6114
Qy 2938 CTCACGGAACATGCGCTGGTGAAGGTTGGGGTGTGGCTGTGAGAGGAGCTGAAGAA 2997
Db 6115 CTCACGGAACATGCGCTGGTGAAGGTTGGGGTGTGGCTGTGAGAGGAGCTGAAGAA 6174
Qy 2998 AAGAAGATGTTCTGACAGACTCGGCTTGAAGGATGTGACCTGATCTGATGCTGCT 3057
Db 6175 AAGAAGATGTTCTGACAGACTCGGCTTGAAGGATGTGACCTGATCTGATGCTGCT 6234
Qy 3058 GAGGCACTCTCTTGGCCCTTCTGCAACCGGCACTGACCAAGTCTAAACCTGTGAG 3117
```

```
Db 6235 GAGGCACTCTCTTGGCCCTTCTGCAACCGGCACTGACCACTGATTAACCTGTGAG 6294
Qy 3118 AATACTTCACTGCTCCCAAGGAATGATGAGCTGTGTTGCGGCTTGTGCTGCTCCAGCTCT 3177
Db 6295 AATACTTCACTGCTCCCAAGGAATGATGAGCTGTGTTGCGGCTTGTGCTGCTCCAGCTCT 6354
Qy 3178 AACTTACAGATTAATTGG 3194
Db 6355 AACTTACAGATTAATTGG 6371

RESULT 9
AX459891
LOCUS Sequence 34 from Patent WO240668.
DEFINITION AX459891
ACCESSION AX459891
VERSION AX459891.1 GI:21725654
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Teschopp, J. and Martinon, F.
TITLES Proteins and dna sequences underlying these proteins used for
creating Inflammation
JOURNAL Patent: WO 0240668-A 34 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source
1..6939
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP13/Py17.cdna"

ORIGIN
Query Match 75.8%; Score 3057.8; DB 6; Length 6939;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;

Qy 1 ATGAGAGAGCAAAATGGCTCACTTTTCCAGCTACGCGGCTGCAATGCTGCTTATGAG 60
Db 3055 ATGAGAGAGCAAAATGGCTCACTTTTCCAGCTACGCGGCTGCAATGCTGCTTATGAG 3114
Qy 61 CTAGACAGAGAAATTTGAGACATTCAGAAATTAATAAGAAATCTTCAGAAATCG 120
Db 3115 CTAGACAGAGAAATTTGAGACATTCAGAAATTAATAAGAAATCTTCAGAAATCG 3174
Qy 121 ACCACATGCTCTATTTCAACAGTTTGAATGAGAAATGCAACGTAATGTCTGGCACTC 180
Db 3175 ACCACATGCTCTATTTCAACAGTTTGAATGAGAAATGCAACGTAATGTCTGGCACTC 3234
Qy 181 CTCTTGATAGATATTATGAGACATCGCTGGCCCTGGGCTAGCTTCATTAAGATTTGAA 240
Db 3235 CTCTTGATAGATATTATGAGACATCGCTGGCCCTGGGCTAGCTTCATTAAGATTTGAA 3294
Qy 241 AACATTAACCTGCGAACCCTCTGAGAGAGCAAGGATGACATGAATAA----- 289
Db 3295 AACATTAACCTGCGAACCCTCTGAGAGAGCAAGGATGACATGAATAAATTTCACCA 3354
Qy 290 ----- 289
Db 3355 GAAGATCTGAAGCAAGATGACTGACCAAGAACCAAGAAATAAGTCAGAAATT 3414
Qy 290 -----AAATTCA 297
Db 3415 TCACAACTGTGCAACAAAGTATGTCACAGCTGACAGAACAAAGAAATAATTTC 3474
Qy 298 CAAGCTATGGAACAAGAAAGGTCACAGACAGAGACAGAGAAACAAGAAATTTCA 357
Db 3475 CAAGCTATGGAACAAGAAAGGTCACAGACAGAGACAGAGAAACAAGAAATTTCA 3534
Qy 358 GCTATGGAACAAGAAAGGTCACAGACAGAGACAGAGAAACAAGAAACAATGAGTGC 417
```

Db 3535 GGTATGAAACAAAGAGTGCACAGACGAGAGACAAAGAAACAAGACATGAGAGTGCAC 3594  
Qy 418 ACATGGAGACTACAAAGATCACTGATGACCAAAATTCGCTGAGAGAGAGATGTAAGTCTGT 477  
Db 3595 ACATGGAGACTACAAAGATCACTGATGACCAAAATTCGCTGAGAGAGAGATGTAAGTCTGT 3654  
Qy 478 AGTTTGAAGAACATGCTGCTGACCTGAGCGGAAATGCAAAACGTTGCTGTCTTTGAT 537  
Db 3655 AGTTTGAAGAACATGCTGCTGACCTGAGCGGAAATGCAAAACGTTGCTGTCTTTGAT 3714  
Qy 538 TCAGACCGGCTGGGAGCTTCCGGCTCGCACGCTGCTGTCGACGAGAAATCAGGAATGGG 597  
Db 3715 TCAGACCGGCTGGGAGCTTCCGGCTCGCACGCTGCTGTCGACGAGAAATCAGGAATGGG 3774  
Qy 598 AAATGGCTCTAGACCAAGAGATCGTGTGTGTGGGCGCAAGGTGACTTACCAAGGGA 657  
Db 3775 AAATGGCTCTAGACCAAGAGATCGTGTGTGTGGGCGCAAGGTGACTTACCAAGGGA 3834  
Qy 658 ATGTTCTCTAGCTCTTCTTCTTCCCTTGAAGAGATGCAAGCGGAAAGAGAGACAT 717  
Db 3835 ATGTTCTCTAGCTCTTCTTCTTCCCTTGAAGAGATGCAAGCGGAAAGAGAGACAT 3894  
Qy 718 GTCAAGAGATTATCTCCAGGAGTGGCCAGACTCCAGAGCTCCGGTGAACGAGATCATG 777  
Db 3895 GTCAAGAGATTATCTCCAGGAGTGGCCAGACTCCAGAGCTCCGGTGAACGAGATCATG 3954  
Qy 778 TCCCGACCAAGAAAGCTGTGTTTCATCATTTGACGGTTTGATGACCTGTGCTCTGCTC 837  
Db 3955 TCCCGACCAAGAAAGCTGTGTTTCATCATTTGACGGTTTGATGACCTGTGCTCTGCTC 4014  
Qy 838 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACAGCTTCCTGACCTCAT 897  
Db 4015 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACAGCTTCCTGACCTCAT 4074  
Qy 898 CCGAGCTGCTGAGGAAGGCTCCGCTCCCTGAGTCCCTTCTGATGCTGACCCGTCAGAGAC 957  
Db 4075 CCGAGCTGCTGAGGAAGGCTCCGCTCCCTGAGTCCCTTCTGATGCTGACCCGTCAGAGAC 4134  
Qy 958 GTGGGACACAGAAAGCTCAAGTCAAGAGTCTGTCTCCCGTTACCTGTTAGTTAGAGA 1017  
Db 4135 GTGGGACACAGAAAGCTCAAGTCAAGAGTCTGTCTCCCGTTACCTGTTAGTTAGAGA 4194  
Qy 1018 ATCTCCGGGAAACAAAGATCCACTTGTCTTGAAGCGGGAATGGTGAATCAGAG 1077  
Db 4195 ATCTCCGGGAAACAAAGATCCACTTGTCTTGAAGCGGGAATGGTGAATCAGAG 4254  
Qy 1078 AACAAAGGTTGCGTGCATCATGAAACAACGTTGAGCTGTCGACCAAGTGCAGAGGCC 1137  
Db 4255 AACAAAGGTTGCGTGCATCATGAAACAACGTTGAGCTGTCGACCAAGTGCAGAGGCC 4314  
Qy 1138 GCGGTGAGCTCTCATCTGCGTGGCCCTGACGCTGACGAGAGAGTGGTGGGAGAGCGTC 1197  
Db 4315 GCGGTGAGCTCTCATCTGCGTGGCCCTGACGCTGACGAGAGAGTGGTGGGAGAGCGTC 4374  
Qy 1198 GCGCCCTTCAACCAACGCTCAACAGGCTGCAAGCGGCTTTTGTGTTTATCATGACTCAC 1257  
Db 4375 GCGCCCTTCAACCAACGCTCAACAGGCTGCAAGCGGCTTTTGTGTTTATCATGACTCAC 4434  
Qy 1258 CCTGAGGCTGCTGCGGCGCTGCTCAATCTGAGAGAAAGATTCTGTAAGCGCTTC 1317  
Db 4435 CCTGAGGCTGCTGCGGCGCTGCTCAATCTGAGAGAAAGATTCTGTAAGCGCTTC 4494  
Qy 1318 TGCCTATGCTGTGAGAGAGTGTGAATAGAAAGTCACTGTTGATGTGACGACTC 1377  
Db 4495 TGCCTATGCTGTGAGAGAGTGTGAATAGAAAGTCACTGTTGATGTGACGACTC 4554  
Qy 1378 ATGCTTCAAGAGACTGGGAGAGTGTGAGCTCGCTGCTCTGTTTCAATGAACATCTTCTC 1437  
Db 4555 ATGCTTCAAGAGACTGGGAGAGTGTGAGCTCGCTGCTCTGTTTCAATGAACATCTTCTC 4614  
Qy 1438 CAGACAGCACTGTGAGAGATCAACCTTCTTCACTGAGTCTCAGAGACTTCTGT 1497

Db 4615 CAGACAGCACTGTGAGAGATCAACCTTCTTCACTGAGTCTCAGAGACTTCTGT 4674  
Qy 1498 GCGCCCTTGTACTAGTGTGAAGGCTGTGAAATTCAGAGCCAGCTCTGCTCTGTAC 1557  
Db 4675 GCGCCCTTGTACTAGTGTGAAGGCTGTGAAATTCAGAGCCAGCTCTGCTCTGTAC 4734  
Qy 1558 GTTGAAGAACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCATATCACTCGCTT 1617  
Db 4735 GTTGAAGAACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCATATCACTCGCTT 4794  
Qy 1618 TGAATGAAGCGTTTCTTGTGCTGTGACGGAAGCTGAAGAGGCACTGAGAGTGC 1677  
Db 4795 TGAATGAAGCGTTTCTTGTGCTGTGACGGAAGCTGAAGAGGCACTGAGAGTGC 4854  
Qy 1678 CTGCTGGGCTGTCCGCTTCCCTGGGGGTGAAGCAAGCTTGTGACTGGGTCTCTG 1737  
Db 4855 CTGCTGGGCTGTCCGCTTCCCTGGGGGTGAAGCAAGCTTGTGACTGGGTCTCTG 4914  
Qy 1738 TTGGGTCAAGACCTTAATGCAACCAAGGAGACCTGAGAGGAGCTTCACTGCTT 1797  
Db 4915 TTGGGTCAAGACCTTAATGCAACCAAGGAGACCTGAGAGGAGCTTCACTGCTT 4974  
Qy 1798 TTGAGACTCAAGACAAAGATTTGTGCTTGGCTTGAATTAACAGCTTCCAAAGATGTG 1857  
Db 4975 TTGAGACTCAAGACAAAGATTTGTGCTTGGCTTGAATTAACAGCTTCCAAAGATGTG 5034  
Qy 1858 CTTCCGATTACCAAGACCTGGAATTGATAGATTTCTTCTGCTTCCAGACTGTCCG 1917  
Db 5035 CTTCCGATTACCAAGACCTGGAATTGATAGATTTCTTCTGCTTCCAGACTGTCCG 5094  
Qy 1918 TATTTGCGGAAATTCGGGTGATGTCAAAGGATCTTCCAAAGATGAGTCCCTGAG 1977  
Db 5095 TATTTGCGGAAATTCGGGTGATGTCAAAGGATCTTCCAAAGATGAGTCCCTGAG 5154  
Qy 1978 GCATGCTCTGTGCTCTCTATGATGCGGATTAAGACCTTCAATTGAGAGAGTGGGA 2037  
Db 5155 GCATGCTCTGTGCTCTCTATGATGCGGATTAAGACCTTCAATTGAGAGAGTGGGA 5214  
Qy 2038 GATTTGCTGCTCATCTTTGGACCCGACCACTGCGGAGCTGAGCTTGGGACGAGC 2097  
Db 5215 GATTTGCTGCTCATCTTTGGACCCGACCACTGCGGAGCTGAGCTTGGGAGCAGC 5274  
Qy 2098 ATCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAAGCATTCGACCTGAG 2157  
Db 5275 ATCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAAGCATTCGACCTGAG 5334  
Qy 2158 ATACAGACCTGATGTTTGAAGAAATGACAGATTAACCTGCTGTGACAGACTCTGAGA 2217  
Db 5335 ATACAGACCTGATGTTTGAAGAAATGACAGATTAACCTGCTGTGACAGACTCTGAGA 5394  
Qy 2218 ATGCTATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGCACCTGAAGAA 2277  
Db 5395 ATGCTATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGCACCTGAAGAA 5454  
Qy 2278 GAGAGTGAAGATGCGGTGAGAGCTTAAACACCCAAATGTTTGTGAGTCTTGG 2337  
Db 5455 GAGAGTGAAGATGCGGTGAGAGCTTAAACACCCAAATGTTTGTGAGTCTTGG 5514  
Qy 2338 AGGCTGAATGCTGAGATTAACCAAGCTGTTTACCTGAAGATTCCTCAATCTTACG 2397  
Db 5515 AGGCTGAATGCTGAGATTAACCAAGCTGTTTACCTGAAGATTCCTCAATCTTACG 5574  
Qy 2398 ACCCTCCCAAGCTGAAATCTCTGAGCTTGAAGGAAACAAAGTGAAGACCAAGGAGTA 2457  
Db 5575 ACCCTCCCAAGCTGAAATCTCTGAGCTTGAAGGAAACAAAGTGAAGACCAAGGAGTA 5634  
Qy 2458 ATGCTCTCAAGTATGCTTGAAGATCTTCAAGTGGCGCTTGAAGACTGATTAAGAG 2517  
Db 5635 ATGCTCTCAAGTATGCTTGAAGATCTTCAAGTGGCGCTTGAAGACTGATTAAGAG 5694  
Qy 2518 GACTGTGACATCAACGAGCGGTTGCAAGATCTGAGCTGCTGCTGACAGAACCGG 2577  
Db 5695 GACTGTGACATCAACGAGCGGTTGCAAGATCTGAGCTGCTGCTGACAGAACCGG 5754



QY 2578 AGCTTGACACACTGTGCTTATCCAAACAGCCCTGGGAAAGAGTGAATCTACTG 2637  
Db 5755 AGCTTGACACACTGTGCTTATCCAAACAGCCCTGGGAAAGAGTGAATCTACTG 5814  
QY 2638 TGTGCATCATGAGGCTTCCCACTGATGCTGCAGAGGCTGATGCTGAATCATGTCAC 2697  
Db 5815 TGTGCATCATGAGGCTTCCCACTGATGCTGCAGAGGCTGATGCTGAATCATGTCAC 5874  
QY 2698 CTGACAGAGGCTGAGTGTGTTTCTTGCACCTGAGGCTTATGAGGTAATCATGCTGACG 2757  
Db 5875 CTGACAGAGGCTGAGTGTGTTTCTTGCACCTGAGGCTTATGAGGTAATCATGCTGACG 5934  
QY 2758 CACCTGAGCCTTAGCATGAACCTCTGGAAGACAATGAGCTTCTGTGCGAGGTC 2817  
Db 5935 CACCTGAGCCTTAGCATGAACCTCTGGAAGACAATGAGCTTCTGTGCGAGGTC 5994  
QY 2818 ATGAGAGAACCATCTTGTCTATCTCCAGGACCTGAGGTTGTAAGTGTCTATCTCACGCC 2877  
Db 5995 ATGAGAGAACCATCTTGTCTATCTCCAGGACCTGAGGTTGTAAGTGTCTATCTCACGCC 6054  
QY 2878 GCGTCTGTGAGAGTCTTCTGCTGTGATCTTCGAGAGACAGACCTGAAGAGCTTGAT 2937  
Db 6055 GCGTCTGTGAGAGTCTTCTGCTGTGATCTTCGAGAGACAGACCTGAAGAGCTTGAT 6114  
QY 2938 CTCACGACCAATGCGCTGGGTGACGCTGGGTTGCTGCGTGGAGGGAAGCA 2997  
Db 6115 CTCACGACCAATGCGCTGGGTGACGCTGGGTTGCTGCGTGGAGGGAAGCA 6174  
QY 2998 AAGAACAGTGTCTGACGAGACTCGGTTGAAGGCACTGACTTCTGATTTGCTGT 3057  
Db 6175 AAGAACAGTGTCTGACGAGACTCGGTTGAAGGCACTGACTTCTGATTTGCTGT 6234  
QY 3058 GAGGACCTCTGCTGGGCTTCTCGCAACGGGCACTGACAGTCTAACTGGTGGAG 3117  
Db 6235 GAGGACCTCTGCTGGGCTTCTCGCAACGGGCACTGACAGTCTAACTGGTGGAG 6294  
QY 3118 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGCGCTTTCCTGTCCACGCTCT 3177  
Db 6295 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGCGCTTTCCTGTCCACGCTCT 6354  
QY 3178 AACTTACAGATTAATGG 3194  
Db 6355 AACTTACAGATTAATGG 6371

RESULT 10  
CQ731113 2753 bp DNA linear PAT 03-FEB-2004  
LOCUS CQ731113  
DEFINITION Sequence 17047 from Patent WO02068579.  
ACCESSION CQ731113  
VERSION CQ731113.1 GI:42306827  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE kites, such as nucleic acid arrays, comprising a majority of  
humanecons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 17047 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Source  
1..2753  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Query Match 63.3%; Score 2553.8; DB 6; Length 2753;  
Best Local Similarity 94.1%; Pred. No. 0;

Matches 2736; Conservative 0; Mismatches 2; Indels 171; Gaps 1;  
QY 404 GACATGAGGTGACACATGGGACTTACAAAGTCAAGTATGACCAAAATTCGCTGAGGAG 463  
Db 1 GACATGAGGTGACACATGGGACTTACAAAGTCAAGTATGACCAAAATTCGCTGAGGAG 60  
QY 464 AGATGTACGTCGTATGTTTGAATAACCTGCTGTGACATGCGCGGAAATGCAACGTTGG 523  
Db 61 AGATGTACGTCGTATGTTTGAATAACCTGCTGTGACATGCGCGGAAATGCAACGTTGG 120  
QY 524 CTGCTGCTTTTGAATTCAGACCGGTGGGCTTCGCGCTTCGACCGGTGTTCTGCACGGA 583  
Db 121 CTGCTGCTTTTGAATTCAGACCGGTGGGCTTCGCGCTTCGACCGGTGTTCTGCACGGA 180  
QY 584 AGTCAGGAATTTGGGAAATCGGCTTACGACGAAGATCGTGTGCTGGGCGCAAGGTG 643  
Db 181 AGTCAGGAATTTGGGAAATCGGCTTACGACGAAGATCGTGTGCTGGGCGCAAGGTG 240  
QY 644 GACTCTACAGGGAATGTTCTCTACGTCCTTCTCTCCCGTTAGAGAGATGACAGCGGA 703  
Db 241 GACTCTACAGGGAATGTTCTCTACGTCCTTCTCTCCCGTTAGAGAGATGACAGCGGA 300  
QY 704 AAGAGAGAGCAGTGTCAAGAGTTCACTTCAGAGAGATGAGCCAGATCTCCAGGCTCGG 763  
Db 301 AAGAGAGAGCAGTGTCAAGAGTTCACTTCAGAGAGATGAGCCAGATCTCCAGGCTCGG 360  
QY 764 TGACGAGATCATATGCCGACCAAGAAAGGCTGTGTTCAATGATGACGTTTCGATGAC 823  
Db 361 TGACGAGATCATATGCCGACCAAGAAAGGCTGTGTTCAATGATGACGTTTCGATGAC 420  
QY 824 TGGGCTCTGTCTCTCAACATGACCAAAAGCTCTGCAAAAGCTGAGGCTGAGAGAGCTC 883  
Db 421 TGGGCTCTGTCTCTCAACATGACCAAAAGCTCTGCAAAAGCTGAGGCTGAGAGAGCTC 480  
QY 884 CGTTCAACCTCATACGAGTGTGTCAGAGAGGTCCTGCTCTGATGCTTCTGATGAG 943  
Db 481 CGTTCAACCTCATACGAGTGTGTCAGAGAGGTCCTGCTCTGATGCTTCTGATGAG 540  
QY 944 TCACGTCAGAGAGTGGGACAGAGAAAGCTCAAGTCAGAGAGTGTGTCCTCCGTTACC 1003  
Db 541 TCACGTCAGAGAGTGGGACAGAGAAAGCTCAAGTCAGAGAGTGTGTCCTCCGTTACC 600  
QY 1004 TGTATGTTAGAGATCTCCGCGGAAACAAAGATCACTTGTCTTGAAGCGGGAATG 1063  
Db 601 TGTATGTTAGAGATCTCCGCGGAAACAAAGATCACTTGTCTTGAAGCGGGAATG 660  
QY 1064 GTGACATCAGAAACACAAAGGTTGCTGATCATGAAACACCGTGAAGTCTCGACC 1123  
Db 661 GTGACATCAGAAACACAAAGGTTGCTGATCATGAAACACCGTGAAGTCTCGACC 720  
QY 1124 AGTCCAGGTCGCGCGTGGGCTCTCATCTGCGTGGGCTGTCAGCTGACGAGAGCTGG 1183  
Db 721 AGTCCAGGTCGCGCGTGGGCTCTCATCTGCGTGGGCTGTCAGCTGACGAGAGCTGG 780  
QY 1184 TGGGGAAGAGGTCGCGGCTTCAACCAACGCTCAGAGGCTGACAGCGCTTTTGTGT 1243  
Db 781 TGGGGAAGAGGTCGCGGCTTCAACCAACGCTCAGAGGCTGACAGCGCTTTTGTGT 840  
QY 1244 TTTATCAGCTCACCCCTCGAGAGGCTGTGCGGCTGTCTCAATCTGAGAGAAAGATTG 1303  
Db 841 TTTATCAGCTCACCCCTCGAGAGGCTGTGCGGCTGTCTCAATCTGAGAGAAAGATTG 900  
QY 1304 TCCGAGAGCGTTTGTGCGATATGCTGTGAGAGGAGTGTGAATGAGAAAGTCAATGTTTG 1363  
Db 901 TCCGAGAGCGTTTGTGCGATATGCTGTGAGAGGAGTGTGAATGAGAAAGTCAATGTTTG 960  
QY 1364 ATGTGACGACCTTCATGATTCAGAGACTCGGAGAGTCTGAGCTTCGCTCTGTTTCA 1423  
Db 961 ATGTGACGACCTTCATGATTCAGAGACTCGGAGAGTCTGAGCTTCGCTCTGTTTCA 1020  
QY 1424 TGAACATCTCTTCCCAAGACGCACTGTGAGAGTACTACACTTCTTCCACTCACTCAGTC 1483  
Db 1021 TGAACATCTCTTCCCAAGACGCACTGTGAGAGTACTACACTTCTTCCACTCACTCAGTC 1080



1484 TCAGAGACTTCTGTGCGGCTTGTACTAGTGTAGAGGCGCTGGAATGAGCCAGCTC 1543  
 1081 TCAGAGACTTCTGTGCGGCTTGTACTAGTGTAGAGGCGCTGGAATGAGCCAGCTC 1140  
 1544 TCTGCGCTCTGTACCTTGAAGAGACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCC 1603  
 1141 TCTGCGCTCTGTACCTTGAAGAGACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCC 1200  
 1604 ATATCACTCGCTTGTGATGAGAGCGTTTCTGTTGGCTCTGTAGCGGAAGCTTAAAG 1663  
 1201 ATATCACTCGCTTGTGATGAGAGCGTTTCTGTTGGCTCTGTAGCGGAAGCTTAAAG 1260  
 1664 GGGCACTGAGAGTCTGTGCGGCTGCTCCGTTCCCTGGGGGTGAAGAGAGCTTCC 1723  
 1261 GGGCACTGAGAGTCTGTGCGGCTGCTCCGTTCCCTGGGGGTGAAGAGAGCTTCC 1320  
 1724 ACTGGGTCTCTGTGAGGTGAGAGAGCTTAAATGAGCCAGAGAGAGAGCTTGAAG 1783  
 1321 ACTGGGTCTCTGTGAGGTGAGAGAGCTTAAATGAGCCAGAGAGAGAGCTTGAAG 1380  
 1784 CTTTCACTGTCTTTTCAAGACTCAAGACAAAGAGTTGTTGCTTGGCATTAACAGCT 1843  
 1381 CTTTCACTGTCTTTTCAAGACTCAAGACAAAGAGTTGTTGCTTGGCATTAACAGCT 1440  
 1844 TCCAGAGAGTGTGCTTCCGATTAAACAGAGCTTGAATGAGCTTCTCTGCTG 1903  
 1441 TCCAGAGAGTGTGCTTCCGATTAAACAGAGCTTGAATGAGCTTCTCTGCTG 1500  
 1904 TCCAGAGAGTGTGCTTCCGATTAAACAGAGCTTGAATGAGCTTCTCTGCTG 1963  
 1501 TCCAGAGAGTGTGCTTCCGATTAAACAGAGCTTGAATGAGCTTCTCTGCTG 1560  
 1964 ATGAGTCTGCTGAGAGAGTCTGTGCTCTGTATGAGTGGAGATTAAGAGCTTCTG 2023  
 1561 ATGAGTCTGCTGAGAGAGTCTGTGCTCTGTATGAGTGGAGATTAAGAGCTTCTG 1620  
 2024 AGAGAGAGTGGAGAAATTTTGTCTGATGCTTGGAGAGAGAGAGAGAGAGAGAG 2083  
 1621 AGAGAGAGTGGAGAAATTTTGTCTGATGCTTGGAGAGAGAGAGAGAGAGAGAG 1680  
 2084 ACCTGGAG 2143  
 1681 ACCTGGAG 1740  
 2144 ATCCAG 2203  
 1741 ATCCAG 1800  
 2204 AGAGAGAGTGGAGAAATTTTGTCTGATGCTTGGAGAGAGAGAGAGAGAGAGAG 2263  
 1801 AGAGAGAGTGGAGAAATTTTGTCTGATGCTTGGAGAGAGAGAGAGAGAGAGAG 1860  
 2264 CCACCTTGAAG 2323  
 1861 CCACCTTGAAG 1920  
 2324 TGTGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2383  
 1921 TGTGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1936  
 2384 CCACCTTGAAG 2443  
 1937 CCACCTTGAAG 1936  
 2444 CAG 2503  
 1937 CAG 1936  
 2504 AGCTGATATCTGAG 2563  
 1937 AGCTGATATCTGAG 1989

2564 TGTGAG 2623  
 1990 TGTGAG 2049  
 2624 GTGTAATCTATCTGTGTGATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2683  
 2050 GTGTAATCTATCTGTGTGATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2109  
 2684 TGAATCAATGAG 2743  
 2110 TGAATCAATGAG 2169  
 2744 ACTGATGAG 2803  
 2170 ACTGATGAG 2229  
 2804 TTTCTGTGAG 2863  
 2230 TTTCTGTGAG 2289  
 2864 GTGATCTACAG 2923  
 2290 GTGATCTACAG 2349  
 2924 TGAAG 2983  
 2350 TGAAG 2409  
 2984 AGGAG 3043  
 2410 AGGAG 2469  
 3044 CTTGATGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3103  
 2470 CTTGATGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2529  
 3104 TAAACCTGTGAG 3163  
 2530 TAAACCTGTGAG 2589  
 3164 CTTGATGATGCTGTGAG 3223  
 2590 CTTGATGATGCTGTGAG 2649  
 3224 TAAAG 3283  
 2650 TAAAG 2709  
 3284 GGCATCTCTTGTGAG 3312  
 2710 GGCATCTCTTGTGAG 2738

RESULT 11  
 AC011470/c 157141 bp DNA linear PRI 15-JUL-2000  
 LOCUS Homo sapiens chromosome 19 clone CTC-490M10, complete sequence.  
 DEFINITION AC011470  
 ACCESSION AC011470.5 GI:9211204  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 157141)  
 DOE Joint Genome Institute and Stanford Human Genome Center.  
 AUTHORS  
 Direct Submission  
 JOURNAL  
 Unpublished  
 2 (bases 1 to 157141)  
 REFERENCE  
 DOE Joint Genome Institute.  
 AUTHORS  
 Direct Submission  
 TITLE  
 Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 157141)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 On Jul 15, 2000 this sequence version replaced gi:7690109.  
 COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.sbgc.stanford.edu  
 Quality: Phrap Quality >=40 99.9% of Sequence;  
 Estimated Total Number of Errors is 0.1.

## FEATURES

source  
 1..157141  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="CTC-490M1.0"

## ORIGIN

Query Match 39.6%; Score 1598.8; DB 9; Length 157141;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 400 CAAGGACATGAGGTCACATGAGGATCAAGAGTCAGTATGACCAAAATCGCTGAG 459
DB 42003 CCAAGACATGAGGTCACATGAGGATCAAGAGTCAGTATGACCAAAATCGCTGAG 41944
QY 460 GAGGAGATGATCGTCTGATGTTTGAACAACCTGCTGCTGACCTGCGGAAATGCAACG 519
DB 41943 GAGGAGATGATCGTCTGATGTTTGAACAACCTGCTGCTGACCTGCGGAAATGCAACG 41884
QY 520 TTGCTGCTGCTTTTGAATTCAGACCGGCTTCCGCTCCGACCGTATGCTGAC 579
DB 41883 TTGCTGCTGCTTTTGAATTCAGACCGGCTTCCGCTCCGACCGTATGCTGAC 41824
QY 580 GAAAGTCAGGAATTTGGAAATCGCTTACGACGAAGATGCTGCTGCTGCGGCA 639
DB 41823 GAAAGTCAGGAATTTGGAAATCGCTTACGACGAAGATGCTGCTGCTGCGGCA 41764
QY 640 GGTGACCTTACGAGGAATGTTCTCTACGCTTTCTTCCCTCCGTTAGAGATGAG 699
DB 41763 GGTGACCTTACGAGGAATGTTCTCTACGCTTTCTTCCCTCCGTTAGAGATGAG 41704
QY 700 CGGAAGAGAGAGAGATGTCACAGATTCATCTCCAGGAGTGCACGACTCCAGGCT 759
DB 41703 CGGAAGAGAGAGAGATGTCACAGATTCATCTCCAGGAGTGCACGACTCCAGGCT 41644
QY 760 CCGGTGACGAGATCATGCTCCGACACAGAAAGGCTGTTGTTCAATCATGACGGTTGAT 819
DB 41643 CCGGTGACGAGATCATGCTCCGACACAGAAAGGCTGTTGTTCAATCATGACGGTTGAT 41584
QY 820 GACTGGGCTTGTCTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGCAG 879
DB 41583 GACTGGGCTTGTCTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGCAG 41524
QY 880 CCTCGCTTACCCCTCATAGAGTCTGCTGAGAAAGCTGCTCCGAGTCTTCCCTG 939
DB 41523 CCTCGCTTACCCCTCATAGAGTCTGCTGAGAAAGCTGCTCCGAGTCTTCCCTG 41464
QY 940 ATCGTACCGTCAAGAGCTGGGACAGAGAGCTCAAGTCAAGAGTCTTCCCGT 999
DB 41463 ATCGTACCGTCAAGAGCTGGGACAGAGAGCTCAAGTCAAGAGTCTTCCCGT 41404
QY 1000 TACCTGTTAGTTAAGAAATCTCCGAGGACCAAAAGATTCATCTGCTTGAAGCGGG 1059
DB 41403 TACCTGTTAGTTAAGAAATCTCCGAGGACCAAAAGATTCATCTGCTTGAAGCGGG 41344
QY 1060 ATTGCTGATCAGCAAGAGCTGGTGGTGGATCAAGAACACCGTGAAGTGTCT 1119
DB 41343 ATTGCTGATCAGCAAGAGCTGGTGGTGGATCAAGAACACCGTGAAGTGTCT 41284

```

```

QY 1120 GACCAATGCCAGTGGCCGCGTGGGCTCTCTCATCTGCTGCTGACCTGACAGTCCAGAGAC 1179
DB 41283 GACCAATGCCAGTGGCCGCGTGGGCTCTCTCATCTGCTGCTGACCTGACAGTCCAGAGAC 41224
QY 1180 GTGGTGGGAGAGCGTGGCCCTTTCAACCAAAAGCTTACAGGCTTGCACGCGCTTTT 1239
DB 41223 GTGGTGGGAGAGCGTGGCCCTTTCAACCAAAAGCTTACAGGCTTGCACGCGCTTTT 41164
QY 1240 GTGTTTATGACGTCACCCCTCCAGAGCGTGGTCCGCGCTGCTCAATCTGAGAAAGA 1299
DB 41163 GTGTTTATGACGTCACCCCTCCAGAGCGTGGTCCGCGCTGCTCAATCTGAGAAAGA 41104
QY 1300 GTTGTCTGAAGCGCTTCTGCGGTATGAGCTGTGAGAGGAGTGTGAATAGAACTGATG 1359
DB 41103 GTTGTCTGAAGCGCTTCTGCGGTATGAGCTGTGAGAGGAGTGTGAATAGAACTGATG 41044
QY 1360 TTTGATGTGACGACCTTATGTTTCAAGAGACTCGGGAGTGTAGCTCCGCTCTGTTT 1419
DB 41043 TTTGATGTGACGACCTTATGTTTCAAGAGACTCGGGAGTGTAGCTCCGCTCTGTTT 40984
QY 1420 CACATGAACATCTTCTCCAGACAGGCACTGTGAGGAGTACTACACCTTCTCACCTC 1479
DB 40983 CACATGAACATCTTCTCCAGACAGGCACTGTGAGGAGTACTACACCTTCTCACCTC 40924
QY 1480 AGTCTCAGAGCTTCTGTGCGCCTTGTACTACGTGTAGAGGCTGGAATGAGCCA 1539
DB 40923 AGTCTCAGAGCTTCTGTGCGCCTTGTACTACGTGTAGAGGCTGGAATGAGCCA 40864
QY 1540 GCTCTTGCCTCTGTGACGTTGAGAGACAAAGAGTTCATGAGCTTTAAACAGGCGGC 1599
DB 40863 GCTCTTGCCTCTGTGACGTTGAGAGACAAAGAGTTCATGAGCTTTAAACAGGCGGC 40804
QY 1600 TTCCATATCCACTGCGTTTGGATGAGAGCTTTCTGTTTGGCCCTGAGAGCAAGCTA 1659
DB 40803 TTCCATATCCACTGCGTTTGGATGAGAGCTTTCTGTTTGGCCCTGAGAGCAAGCTA 40744
QY 1660 AGAGAGCCACTGAGAGTCTGCTGAGGCTGCTCCGCTTCCCTGAGGAGTGAAGCAGAAGCTT 1719
DB 40743 AGAGAGCCACTGAGAGTCTGCTGAGGCTGCTCCGCTTCCCTGAGGAGTGAAGCAGAAGCTT 40684
QY 1720 CTGACCTGGGCTCTCTGTTGGTTCAGACAGCTTAATGCAACACCCAGAGAGACCTTG 1779
DB 40683 CTGACCTGGGCTCTCTGTTGGTTCAGACAGCTTAATGCAACACCCAGAGAGACCTTG 40624
QY 1780 GAGGCTTCCACTGCTCTTTTTCAGAGCTCAAGACAAAGATTTGTTGCTTGGCATTTAAAC 1839
DB 40623 GAGGCTTCCACTGCTCTTTTTCAGAGCTCAAGACAAAGATTTGTTGCTTGGCATTTAAAC 40564
QY 1840 AGCTTCCAAAGAGTGGGCTTCCGATTAAACCAAGACCTGGAATTGATAGCATCTTCTTC 1899
DB 40563 AGCTTCCAAAGAGTGGGCTTCCGATTAAACCAAGACCTGGAATTGATAGCATCTTCTTC 40504
QY 1900 TGCTTCCAGACCTGCTCCGATTTTGGGAAATTTGGGTGATGTCAAAGGATCTTCCCA 1959
DB 40503 TGCTTCCAGACCTGCTCCGATTTTGGGAAATTTGGGTGATGTCAAAGGATCTTCCCA 40444
QY 1960 AGAGATGAGTCCGCTGAGGAGCATGCTGTGGTCCCTCATGAG 2001
DB 40443 AGAGATGAGTCCGCTGAGGAGCATGCTGTGGTCCCTCATGAG 40402

```

## RESULT 12

AC024580 193609 bp DNA linear PRI 21-DEC-2001  
 LOCUS Homo sapiens chromosome 19 clone CTD-2621117, complete sequence.  
 DEFINITION AC024580  
 ACCESSION AC024580.6 GI:17975240  
 VERSION HTG.

KEYWORDS  
 SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 193609)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 193609)  
AUTHORS DOE Joint Genome Institute.  
JOURNAL Direct Submission  
SUBMITTED (29-FEB-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 193609)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
SUBMITTED (13-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 193609)  
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
JOURNAL Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Dec 21, 2001 this sequence version replaced gi:16905144.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-sbgc.stanford.edu  
Quality: Phrap Quality >=40.99.4% of Sequence;  
Estimated Total Number of Errors is 0.9.  
FEATURES  
source  
1..193609  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-262117"  
ORIGIN  
Query Match 39.6%; Score 1598.8; DB 9; Length 193609;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
400 CAGAGCATGAGGTGACATGGAATGAGAGTACAGTATGACCAATTCGCTGAG 459  
22255 CAGAGCATGAGGTGACATGGAATGAGAGTACAGTATGACCAATTCGCTGAG 22314  
460 GAGGAGATGATGCTGATGTTTGAACACATGCTGCTGACGCGCGGAAATGCAACG 519  
22315 GAGGAGATGATGCTGATGTTTGAACACATGCTGCTGACGCGCGGAAATGCAACG 22374  
520 TTGGCTGCTGCTTGTGATTCAGACCGGTGGGCTTCGCGCTCGACGCGTGTCTGCAC 579  
22375 TTGGCTGCTGCTTGTGATTCAGACCGGTGGGCTTCGCGCTCGACGCGTGTCTGCAC 22434  
580 GGAAGATCAGAAATGGGAATCGGCTTAGCCAGAAAGATGCTGCTGCTGGGCGCA 639  
22435 GGAAGATCAGAAATGGGAATCGGCTTAGCCAGAAAGATGCTGCTGCTGGGCGCA 22494  
640 GGTGACTTACCAAGGAATGTTTCTCCTACGCTTCTTCTCCCGTTAGAGATGCGAG 699  
22495 GGTGACTTACCAAGGAATGTTTCTCCTACGCTTCTTCTCCCGTTAGAGATGCGAG 22554  
700 CGGAGAGAGAGAGAGTCAAGAGTTCATCTCAGGAGATGAGCAGACTCCAGGCT 759  
22555 CGGAGAGAGAGAGAGTTCATCTCAGGAGATGAGCAGACTCCAGGCT 22614  
760 CCGGTGACGAGATCATGTCCGACCAAGAGGCTGTGTTTCATCATTTGACGGTTCGAT 819  
22615 CCGGTGACGAGATCATGTCCGACCAAGAGGCTGTGTTTCATCATTTGACGGTTCGAT 22674  
820 GACCTGGGCTCTGTCTCAACAATGACCAAAAGCTTCGAAAGAGTGGGCTGAGAGCAG 879  
22675 GACCTGGGCTCTGTCTCAACAATGACCAAAAGCTTCGAAAGAGTGGGCTGAGAGCAG 22734  
880 CTTCCGTTCAACCTCATAGCAGATGCTGCTGAGGAAGGCTCTGCTCCCTGAGTCTTCCG 939  
22735 CTTCCGTTCAACCTCATAGCAGATGCTGCTGAGGAAGGCTCTGCTCCCTGAGTCTTCCG 22794

940 ATTCGACCCGTGAGAGAGTGGGCA CAGAAAGCTCAAGTCAAGAGTGTGCTCCCGT 999  
22795 ATTCGACCCGTGAGAGAGTGGGCA CAGAAAGCTCAAGTCAAGAGTGTGCTCCCGT 22854  
1000 TACCTGTAGTTAGAGATCTCCGGGAA CAAAGATTCACCTTCTCTTGAAGCGCGG 1059  
22855 TACCTGTAGTTAGAGATCTCCGGGAA CAAAGATTCACCTTCTCTTGAAGCGCGG 22914  
1060 ATTGCTGAGCATCAGAAACACAGAGG TTTGGCTGCTGATCATGACCAACCTGAGCTGCTC 1119  
22915 ATTGCTGAGCATCAGAAACACAGAGG TTTGGCTGCTGATCATGACCAACCTGAGCTGCTC 22974  
1120 GACCAATCCAGAGTCCCGCGGCTCTCT CACTGCTGCTGCTGCTGCTGCTGAGGAC 1179  
22975 GACCAATCCAGAGTCCCGCGGCTCTCT CACTGCTGCTGCTGCTGCTGCTGAGGAC 23034  
1180 GTGCTGGGGAGAGAGTGGCGCCCTTCA ACACAGCTCACAGGCTGACACGCGCTTCT 1239  
23035 GTGCTGGGGAGAGAGTGGCGCCCTTCA ACACAGCTCACAGGCTGACACGCGCTTCT 23094  
1240 GTTTCATCAGCTCAGCCCTTCGAGGCG TGTCCGCGCTGTCTCATCTGAGGAAAGA 1299  
23095 GTTTCATCAGCTCAGCCCTTCGAGGCG TGTCCGCGCTGTCTCATCTGAGGAAAGA 23154  
1300 GTTTCATCAGAGCGCTTCTGCGCTGAT TGGCTGAGAGGAGAGTGTGGAATAGAACTCAGTG 1359  
23155 GTTTCATCAGAGCGCTTCTGCGCTGAT TGGCTGAGAGGAGAGTGTGGAATAGAACTCAGTG 23214  
1360 TTTGATGCTGAGAGCCTCATGAGTTCAAG AACTCGGGAGTCTGAGCTCGTCTCTGTTT 1419  
23215 TTTGATGCTGAGAGCCTCATGAGTTCAAG AACTCGGGAGTCTGAGCTCGTCTCTGTTT 23274  
1420 CACATGAACATCTCTTCTCCAGACAGCA CTTGTGAGAGATCTACACCTTCTTCACCTC 1479  
23275 CACATGAACATCTCTTCTCCAGACAGCA CTTGTGAGAGATCTACACCTTCTTCACCTC 23334  
1480 AGTCTTCAGAGATTTCTGCTGCGCGCT TGTACTACGTTTAGAGGCGCTGGAATTCAGGCA 1539  
23335 AGTCTTCAGAGATTTCTGCTGCGCGCT TGTACTACGTTTAGAGGCGCTGGAATTCAGGCA 23394  
1540 GCTCTCTGCGCTCTGTAGTTGAGAAACA AAGAGGCTCATGAGCTTAAACAGGCAAGC 1599  
23395 GCTCTCTGCGCTCTGTAGTTGAGAAACA AAGAGGCTCATGAGCTTAAACAGGCAAGC 23454  
1600 TTCCATATCACTGCTTGTGATGAGAGGCT TTTGCTTGTGCTCTGAGGCAAGCATA 1659  
23455 TTCCATATCACTGCTTGTGATGAGAGGCT TTTGCTTGTGCTCTGAGGCAAGCATA 23514  
1660 AGAGGCACTGAGAGTCTGCTGAGGCTGT CCGTCCCTGCGGGGTGAGCAAGCTT 1719  
23515 AGAGGCACTGAGAGTCTGCTGAGGCTGT CCGTCCCTGCGGGGTGAGCAAGCTT 23574  
1720 CTGCACTGGGCTCTCTGTGGGTGAGGAGCT AATGCAACACCCGAGAGACACCTG 1779  
23575 CTGCACTGGGCTCTCTGTGGGTGAGGAGCT AATGCAACACCCGAGAGACACCTG 23634  
1780 GACGCTTCACTGCTTGTGAGAGCTCAAGA CAAAGTGTGCTGCTGAGCATTAAC 1839  
23635 GACGCTTCACTGCTTGTGAGAGCTCAAGA CAAAGTGTGCTGCTGAGCATTAAC 23694  
1840 AGCTTCAAGAAAGTGTGCTTCCGATTAA CCAAGACTTGAATGACATCTTCTTC 1899  
23695 AGCTTCAAGAAAGTGTGCTTCCGATTAA CCAAGACTTGAATGACATCTTCTTC 23754  
1900 TGCCTTCAGACCTGTGCTGCTTATTTGGG AATTTGGGGGTGAGTGTCAAGGAGATCTTCCA 1959  
23755 TGCCTTCAGACCTGTGCTGCTTATTTGGG AATTTGGGGGTGAGTGTCAAGGAGATCTTCCA 23814  
1960 AGAGATGAGTCCGCTGAGAGCATGCTGTG TGTGCTCTATGAG 2001  
23815 AGAGATGAGTCCGCTGAGAGCATGCTGTG TGTGCTCTATGAG 23856

RESULT 13  
LOCUS AY721594 3545 bp mRNA linear MAM 15-SEP-2004  
DEFINITION Bos taurus maternal antigen that embryo require mRNA, complete cds.  
ACCESSION AY721594  
VERSION AY721594.1 GI:51980124  
KEYWORDS  
SOURCE .  
ORGANISM Bos taurus (cow)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 3545)  
Penneker, S., Uzbekova, S., Perreau, C., Papillier, P., Merrillod, P. and Dalbès-Tran, R.  
Spatio-Temporal Expression of the Germ Cell Marker Genes MATER, ZAR1, GDF9, BMP15, and VASA in Adult Bovine Tissues, Oocytes, and Preimplantation Embryos  
Biol. Reprod. (2004) in press  
2 (bases 1 to 3545)  
Penneker, S., Uzbekova, S. and Dalbès-Tran, R.  
Direct Submission  
Submitted (12-AUG-2004) Physiologie de la Reproduction et des Comportements, Institut National de la Recherche Agronomique, Nouzilly 37380, France  
LOCATION/Qualifiers  
1..3545  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
1..3297  
/note="MATER; NALP5"  
/codon\_start=1  
/product="maternal antigen that embryo require"  
/protein\_id="AAU20764.1"  
/db\_xref="GI:51980125"  
/translation="\"REAKIAPLSNYGLQMFQOLGKEEPQFALLKEHASESAACS  
FPLVOVDRADESLALHEHCRASIAKSTDIPEKMSIALSMADDEKYLIAE  
ISEDSAPKTDGQSGMEKPGPREDDPSRRIHYMTPTSTRDTPRPEPSECP  
DAHALSGAFNDPSSGFRPLTVLHGPVGVKSSIARLLIFMAQGDLYKLFPSYFL  
LRARDLQGSRTSFAELISKEMPDAPVEVLSQPELLIVDGLLELELTFPDQS  
SLADWAEROPAPLVASHLKLKVLPECALITVQDAIQRLQALRLSPRLWVGLS  
VENRMQALLGGGCKRKCAMHAGADHOEYLDKQVAVVVALVEALELGEPEGKL  
PVGHITLGLYATFVFORLAPKDAQWRLSGEBREGALGLRIADGYWNAKFPDGD  
DLVGHGLOGPELSALQASILLPOGHCGRGAFSLISQEPFALLPYLRVEDDGG  
YPLFPQSTKSLTELRLHIDLNVQVMKRFGLVSKVEYRALLETLLGCPVPAKQQL  
LHWICLVGQHPAAASPDLBAFYCLFEAODEFVRLALNGFQEWQLNPMDLTVS  
SFCLRRQHLRRKRLDVAGTPRDEFAEMSGAPQGLKXTIDHEVEDLCSYLSTHPL  
RQDLDSGVLSKEAMKTLCVLRQPAKCIOMLIFKGAVTGELRLHMTLLINNTIR  
LDLNGCRLEEDVOTACALHHPQCALSLDLRGLPASCRLSIQVLAATSGLSKSL  
SLTGKLVADQVTKSLCDALKYTPCTIOGLIISGCGLTAAQODLALILEMOGLTHLS  
LSGRLSGSKMSLUCRAVKLSSCGIQLKALNACSLDVAGCFPLAFALMGNHHTLSL  
SMNPLEDGPMNLCEVMEPSCPLKDLVNCRLTASCKSLSVITRPSRLSDLA  
ANALGDEGIALCAGLKQKNTLTRLGLRAGVLTSEGCALNALITCSHSLANIMRN  
DLGPGMTLLCSAFMHPTSNLQITGLMREQYPARVRLLEQVORLKHVIVISDAWYTE  
EEEDGPCWRI\"  
1..291  
/note="Region: pyrin/DAPIN domain"  
538..1047  
/note="Region: NACHT domain"  
2035..2118  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
2203..2286  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
2290..2373  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
2374..2457  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"

misc\_feature 2461..2544  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2545..2628  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2632..2715  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2716..2799  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2803..2886  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2887..2970  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2971..3054  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 3055..3138  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
3'UTR  
polyA\_signal  
polyA\_site 3545  
ORIGIN  
Query Match 38.8%; Score 1566; DB 4; Length 3545;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 2271; Conservative 0; Mismatches 1005; Indels 42; Gaps 6;  
QY 1 ATGAGAGGAGCAATATGCTCACTTTCCAGCTACGAGGCGTGCATGTGCTCTATGAG 60  
DB 1 ATGAGAGAGGCAAAATAGCCCTCTATTCATTAAGACTGCACTGTGTTTGAGCAG 60  
QY 61 CTGAGCAAGAGAAATTTACAGACATTCAGAGAAATTAAGAAATCTTCAGAAATCG 120  
DB 61 CTGGCAAGAGAGATTCACAGATTTAAGCCCTCTGAAAGAGACATGCTCAGAAATCA 120  
QY 121 ACCAATGCTCTATTTCCACAGTTGAAATCGAAGATCGAAAGTCTGGCACTC 180  
DB 121 GCACAGTGCTCTTCCTCGCTGGTCCAGGTGACAGGCTGATGCGAGTCCCTGCCCTCC 180  
QY 181 CTCTTGATGATGATTAATGAGCATTCCTGGCCCTGGGCTAGTCCATTAAGATCTTTGAA 240  
DB 181 CTCTGATGATGATTAATGAGCATTCCTGGCCCTGGGCTAGTCCATTAAGATCTTTGAA 240  
QY 181 CTCTGATGATGATTAATGAGCATTCCTGGCCCTGGGCTAGTCCATTAAGATCTTTGAA 240  
DB 181 CTCTGATGATGATTAATGAGCATTCCTGGCCCTGGGCTAGTCCATTAAGATCTTTGAA 240  
QY 241 AACATGAACCTGCGAACCCTCTCGAAGAGGACGAGATGATGAAATTTCAAA 300  
DB 241 AACATGAACCTGCGAACCCTCTCGAAGAGGACGAGATGATGAAATTTCAAA 300  
QY 241 AACATGAACCTGCGAACCCTCTCGAAGAGGACGAGATGATGAAATTTCAAA 300  
DB 241 AACATGAACCTGCGAACCCTCTCGAAGAGGACGAGATGATGAAATTTCAAA 300  
QY 301 GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAATTTCAAAAGCT 360  
DB 301 GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAATTTCAAAAGCT 360  
QY 301 GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAATTTCAAAAGCT 360  
DB 301 GCTATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAATTTCAAAAGCT 360  
QY 361 ATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAATTTCAAAAGCT 420  
DB 361 ATGGAACAAGAGGTGCCACAGCAGCAGACAGACAGAAATTTCAAAAGCT 420  
QY 336 --TACCAAGGTCCAAGATGAAGAAAGTGCAGGCCCCAGAGAGACCCACAGAGACTCG 393  
DB 336 --TACCAAGGTCCAAGATGAAGAAAGTGCAGGCCCCAGAGAGACCCACAGAGACTCG 393  
QY 421 TGGAGCTACAGAGCTCAGTATGACCAAAATTCCTGAGAGAGAGATGTACGTGTA 480  
DB 421 TGGAGCTACAGAGCTCAGTATGACCAAAATTCCTGAGAGAGAGATGTACGTGTA 480  
QY 394 CGGACCTACAGAGATCCAGCTATACAGACATTTCTCACACAGGCTGACACACCCACAGGC 453  
DB 394 CGGACCTACAGAGATCCAGCTATACAGACATTTCTCACACAGGCTGACACACCCACAGGC 453  
QY 481 TTGAAAACACTGCTGCTGATGAGCGCGAAATGCAAAAGTTGAGTGTGCTTTGATTCA 540  
DB 481 TTGAAAACACTGCTGCTGATGAGCGCGAAATGCAAAAGTTGAGTGTGCTTTGATTCA 540  
QY 454 TTGAAAAGATTGCTGAGATGTCCGGAACCGCATGCGGATGCGGAGGCTTTAAACCA 513  
DB 454 TTGAAAAGATTGCTGAGATGTCCGGAACCGCATGCGGATGCGGAGGCTTTAAACCA 513  
QY 541 GACCGGT---GGGCTTCGGGCTTCGACAGGTGTTCTGACCGGAAATCAGAAATTCGGG 597  
DB 541 GACCGGT---GGGCTTCGGGCTTCGACAGGTGTTCTGACCGGAAATCAGAAATTCGGG 597  
QY 514 GACCCCTCGGGGGGCTTCGGGCTTCGACAGGTGTTCTGACCGGAAATTCGGG 573  
DB 514 GACCCCTCGGGGGGCTTCGGGCTTCGACAGGTGTTCTGACCGGAAATTCGGG 573

598 AATCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGAAGGTGAATCTACAGAGA 657  
574 AGTCTCGCTGGCCCGGAGGCTCTACTGTCTGGGCGGAGGCACTCTACAGAGG 633  
658 ATGTTCTCTAGCTCTTCTTCTCTCCCGTTAGAGATGCAAGGAGAGAGAGAGT 717  
634 TTGTTCTCTAGCTCTTCTCTCTCCGCGCAGAGACTTCAGAGGGTTCAGAGGAG 693  
718 GTCAAGATTTATCTCTCAGAGAGTGGCCAGACTCCAGAGCTCCGCTGACGAGATCAG 777  
694 TTGCGGAGCTCAATCTCAGAGAGTGGCCGAGAGCCCGGTCCCGGTAGAGAGGTCG 753  
778 TCCCGACAGAAAGCTGTGTGTCTCATTTGACGAGTTTGAGTGAACCTGG---GCTGTGTC 834  
754 TCCGAGCTCTGAAAGGCTCTTGATGTGTGATGATGACTGAGAGGCTGAGGCTCACCTTC 813  
835 CTCAACATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAGAGGAGGCTCCGTTCACTTC 894  
814 AAGAGACAGAGACTCCAGCTCTCGCTGCTGAGGCGGAGAGGAGGAGCCCGCCGCTGTCG 873  
895 ATACGAGTCTGTGAGAGAGTCTGTCTCTGAGTCTTCTGATGCTCACCTGACAG 954  
874 GCGCAGGCTGTGAGAGAGTCTGTCTCCGAGTGGCGCTCTCTCTCACTCCGTCAG 933  
955 GACGTGGGACAGAGAGCTCAAGTGAAGTGTGTCTCCCGTTACCTGTATGTTAGA 1014  
934 GACGCGGAGTTGAGAGGCTCCAGGCTCTCGCTCTCCGTTACTCTATGAGTGGG 993  
1015 GGAATCTCCGAGGAGAACAAAGATTCACCTTGTCTCTTGAAGCGGAGATTGTGAGATCAG 1074  
994 GGCCTCTCAGTGGAGAAACAGAGATGAGTGTGCTCTCGGCGGAGAGAGAGCTGCGTCTG 1053  
1075 AAGACACAGAGGTTGTGTGCTGATCATGAACACCTGAGCTGTGACCACTGAGTCCAGGTG 1134  
1054 AAGACATGCGCTGTGACCGCGGCGGCGGACACAGAGAGTGTGACAGTGTGAGGTG 1113  
1135 CCGCGCGTGGGCTCTCATCTGCGTGGGCTGTGAGCTGACAGAGCTGTGGGAGGAGC 1194  
1114 CCGGTGTGTGGCTGTGAGTGGTGTGAGGCTCTGAACTGACAGAGGAGAGCGGAGAGAGG 1173  
1195 GTGCGCCCTCTCAACCAACGCTCAAGGCTGACAGCGCTTTGTGTATCATGACTC 1254  
1174 CTTCCTGCTCCCGGCGACACCTCAAGGCTTGTAGCCACCTGTGTCTCAGCGGCTG 1233  
1255 ACCCTCGAGGCGTGTGCTGCGGCTGTCTCAATCTGAGAGAAAGATTGCTGAAGCGC 1314  
1234 GCTCCCAAAGATGACAGGCTGCGGCGCTGAGCGGAGAGAGAGCGGCTCCCTGAAGGCGC 1293  
1315 TTCTGCGCTATAGCTGTGAGAGAGTGTGAGATAGGAATCACTGTGTGTATGTGACAC 1374  
1294 TTGTGCGGCTGTGACAGCGGAGCGGCTGTGAAACGCGAGATTGCTGTGTGTGACGCGACAC 1353  
1375 CTCTAGTGTCAAGACTCGGAGAGTCTGAGCTCCGCTCTGTGTTCATCAATCAATCTT 1434  
1354 CTGGGCTCTCAAGGCTGACAGGAGCCGAGCTCTCGGCTCTGACAGAGGCGAGCATCTT 1413  
1435 CTCCAGAGACGACCTGTGAGAGATCTACACTTCTTCACTGACTCTCAGAGACTTC 1494  
1414 CTCCCTGAGCGGCTCACTGCGAGAGGAGCCAGCATTTCTCCACTGAGCTCTCAGAGGTTTC 1473  
1495 TGTGCGCTGTGTATCTACTGTGTAGAGGCTGTGAGAAATGAGCAGCTCTGTGCTCTG 1554  
1474 TTGTGCGCTTGTGTATCTCTGTGAGGCTGTGAGGAGAGCGGAGAGGCTTACCGCTG 1533  
1555 TACGTTGAGAGACAAAGAGTCAATGAGCTTAAACAGGAGGCTTCACTATCACTG 1614  
1534 TTCCCGCAGAGACAAAGAGTCTGACAGAGCTCAAGCATATTAAGCTTCAAGCTTCAAGCTG 1593  
1615 CTCTGATGAGAGCTTCTTGTGTGTGCTGTGAGGAGAGAGTGAAGAGGCACTGAG 1674  
1594 GTCCAGATGAGAGATTTCTATTTGGCTTGTGAGAGAGAGTGAATGCGGCGCTGAG 1653  
1675 GTCTGTGAGGCTGTGCTGCTTCCCTCGGAGGAGTGAAGAGAGCTTGTGACTGAGTCTCT 1734

1654 ACCCTCTAGGCTGTGCTCCGTTGCGCCCGGTGCTAAGCAGCACTTGTGACTGATCTG 1713  
1735 CTGTGGGTGAGAGCTTAATGCTCAACCCAGAGAGACCTCTGAGGCTTCCACTGT 1794  
1714 CTGTGGGTGAGAGCTTGTGCTCCGCTGCTCCAGACTTGTCTGAGGCTTCTACTG 1773  
1795 CTTTTCAGACTCAAGAGACAAAGTGTGTGCTTGGCTTGAATTAAGAGCTTCCAGAGT 1854  
1774 CTTTTCAGAGCTCAGAGAGAGAGTGTGCTTGGCTTGAAGAGCTTCCAGAGT 1833  
1855 TGGCTTCCGATTAAACAGAGACTGATGATCTTCTGCTGCTCCAGAGCTGT 1914  
1834 TGGCTGAGATTAAACCGGCAATGAGCTTAAAGGTGTCTTCTGTCTCCGCGCTG 1893  
1915 CCGTATTTGCGGAAATTTGCGGTGATGCAAGGAGATCTTCCAGAGATGATGCT 1974  
1894 CAGCATTTAAGGAAAGTTGCACTGATGTCAAGAGGAG---TTCAGATGATGATGCT 1950  
1975 GAGGATGTCTGTGTCTCTTATGATGTGGGATTAAGACCTCATTAAGAGAGTGG 2034  
1951 GAGGATGTCTGGGCTCTCCAGAGGCTGAGATCAAGAGCC---TTCAGATGATGATGCT 2007  
2035 GAAAGTTTCTGCTCATGCTTGGGACCAACCCAGCTGCGGAGCTGAGCTGGGAGC 2094  
2008 GAAAGCTTGTCTCTGCTCTGCTGACAGAGCCAACTTACAGAGCTGAGCTGAGTGG 2067  
2095 AGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCGAAGCTGAGGATCCACTG 2154  
2068 AGCTCTTGAAGAGAGAGGCTATGAAGACCTGTGTGCAAGCTGCGGAGCAGCTGT 2127  
2155 AAGATACAGACCTCTGATGTTTGAATGCAAGATTAACCTGTGTGTGAGAGCTTGG 2214  
2128 AAAATACAGATCTGATTTTAAGAGGCGCGGCTTACCCCTGTGTGCTGCTGAG 2187  
2215 AAGATGTGATGCGCAACCGTAACTTAAGATCTCTCACTTGGAGAGGACCCAGCTGAAG 2274  
2188 ATGACTGTGATTTATCAACCGAATCAACGCGCTGAGACTGACGAGCTGCGCTGAG 2247  
2275 GAAAGAGATGTAAGAGTGGCTGTGAGAGCTTAAACACCCCAAAATGTTTGTGAGTCT 2334  
2248 GAGAGAGAGCTGTGACAGCGGCGTGGAGGCGCTCAAGGACCCAGAGCTGAGTGT 2307  
2335 TTGAGCTGATGTGTGTGATGACCAATGCTGTTAACCTGAGATCTCCAAATCTT 2394  
2308 CTGAGCTTGAATGCTGTGTGATTAACCCAGCTCTTGTGAGAGATCTCCAAATGCTT 2367  
2395 ACGACTCCCGAGCTGAATCTGTGAGCTGTGAGAGAAACAGGTGACAGAGGAG 2454  
2368 GTTACGTCCGAGCTGAATCTGTGAGCTTACGAGCTTACGAGAAATGAAGTGGAGGAG 2427  
2455 GTATGCTCTCAAGATGCTGTGAGAGTCTTCCAGTGGCTGCTGAGAGCTGATCTG 2514  
2428 GTGAAGTCTCTGTGAGAGCTTGAAGTCAACCTTGCACCTTGCAGAGAGTATCTG 2487  
2515 GAGAGCTGTGATCAAGAGCAAGGCTTGCAGAGTGTGAGCTGAGCTGTGAGCAAC 2574  
2488 GAGAGCTGTGATCAAGAGCAAGGCTTGCAGAGTGTGAGCTGAGCTGTGAGCAAC 2547  
2575 CGAGCTTGAACACCTGTGTCTATCAACAGAGCTGTGAGAGAGAGAGTGAATCTA 2634  
2548 CAGAGCTTGAACACCTGTGTCTGTGAGAGTGAAGTGTGAGAGAGAGTGAATCTA 2607  
2635 CTGTGTGATCAAGAGCTTCCCACTGATGTGAGAGGCTGATGAGTGAATGAGTGC 2694  
2608 CTGTGTGAGAGCTGTGAGAGCTTCAAGCTGTGTGTGAGAGAGTGTGAGTGAATGAGTGC 2667  
2695 CACTGAGACAGGCTGT 2754  
2668 AGCTGTGATGAGT 2727  
2755 ACGACTGAGCTTGAAGTGAACCTGTGAGAGAGTGTGAGAGTGTGTGTGTGTGTGTGT 2814



Db 2728 ACGCACTGAGCTCAGCATGAAACCCCTGAGAGATCCCGAGATGACCTTCTGTGCCAG 2787  
Qy 2815 GTCATGAGAGAAACATCTTGTATCTCCAGAGCCGTGAGTTGGTAAAGTGCATCTAC 2874  
Db 2788 GTCATGAGAGAGCCGTCTCGCCCTCTCCGGAGCTGGAATTGGTAACTTGCCTGACC 2847  
Qy 2875 GCCCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGACACACTGAAAGCCCTG 2934  
Db 2848 GCCTGTGTGAGAGAGTCTGTCCAACTGATCAGAGAGAGTCCGCGCTGAGAGCCCTG 2907  
Qy 2935 GATCTCAGAGACATAGCCCTGTGGTGAAGCGTGGGTTGCTGCGCTGTGAGAGGATCTGAAG 2994  
Db 2908 GATCTGCGCCGCAAGCCCTGTGGTATGAGGGGATTTGCGCGCTGTGAGAGGATCTGAAG 2967  
Qy 2995 CAAAGAAACAGTGTCTGACGAGACTCGGGGTGGAAGGATGAGCTGACCTTGTATTTGC 3054  
Db 2968 CAGAGAAACA--CCCTGACGAGCTCGGGCTGAGAGCGGTGTGACTGACTCTTAAAGCC 3024  
Qy 3055 TGTGAGGACATCTCTCTGTCCTTCTGCAACCGGATCTGACCGATCTTAAACCTGGTG 3114  
Db 3025 TGCAGAGCCGTGTGCGCGCCCTGACCTGAGCGGCCGCACTGGCCAGCTGAACCTGATG 3084  
Qy 3115 CAGATTAATCTTCAATGCCAAGAGATGATGAAGCTGTGTGGCTTTGCTGTGCCACG 3174  
Db 3085 CGCAACGACCTCGGCCCCAGAGAGATGACGAGCTGTGTCTTGCCTTCAATGATCCCAAC 3144  
Qy 3175 TCTAATCTACAGATTAATGGCGCTGTGGAATGGAGTACCTGTGCAATAAGAGAGCTG 3234  
Db 3145 TCTAATCTACAGACCATCGGCTGTGGAAGAGCAATACCGGCCGAGTAAAGAGGCTG 3204  
Qy 3235 CTGAGAGAGTGCAGCTACTCAAGCCCGAGTGTGATTAATGACGATGTTGGCATTTT 3294  
Db 3205 CTGAGAGAGTGCAGAGACTGAAGGCCCAAGTGTGATGAGAGCCGTGGTGAACGAG 3264  
Qy 3295 GATGAAGTGAACCGACAC 3312  
Db 3265 GAGGAGAGAGAGCGGCC 3282

RESULT 14  
LOCUS AY329487 3405 bp mRNA linear ROD 18-DEC-2003  
DEFINITION Mus musculus strain CBA/J MATR protein isoform-E (Mater) mRNA,  
complete cds; alternatively spliced.  
ACCESSION AY329487  
VERSION AY329487.1 GI:37595472  
KEYWORDS  
SOURCE  
ORGANISM Mus musculus (house mouse)  
MUS musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 3405)  
TITLE Identification and characterization of alternative splice variants  
for murine Mater  
UNPUBLISHED  
JOURNAL 2 (bases 1 to 3405)  
AUTHORS Cheng,H., Huang,Z., Zhang,Y., Liu,H., Teuscher,C. and Ma,R.Z.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUN-2003) Human and Animal Genetics Division,  
Institute of Genetics and Developmental Biology, Chinese Academy of  
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China  
FEATURES  
SOURCE  
1..3405  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CBA/J"  
/db\_xref="taxon:10090"  
1..3405  
/gene="Mater"  
9..3296  
/gene="Mater"  
/note="alternatively spliced"

/codon\_start=1  
/product="Mater protein isoform-E"  
/protein\_id="AA094610.1"  
/db\_xref="GI:37595473"  
/translation="MATERPDKSKILIKARGLSEKSEKSTMSSEVNSKILIDSGSEVQAS  
QEQTSSTWPEKDSKILIKARGLSEKSEKSTMSSEVNSKILIDSGSEVQAS  
ERKMTSPENDSKSIQKQDQPEQEQTSNGNDILKQKAVILAKPTSDVLDHDSBML  
LSDAFKPYOKTFPQHTIILHGRPEVGSALARSIVLQMAQSKLEOKRSFVIFSVREI  
KMTKSSLAOLIAKECPDSMDPVTKIMSOPERLLFVLDGDDMSVLOQDDMTSRDW  
KDBPIYILVLSILKALLPQSLITITVETLEKIMSVSPILYILVEGASARSG  
LVLENISNESDRIVFHSILBNHOLFQCCAPSCSYVCEALQJQKLGKRCCTPCOT  
LTGSLVATLVPHQLTKRPSOALSQEQITLVGICMAAAGVMTSRVFNDDLKNS  
LKSESEIATLPHMNLQVGNHSDCYFSLSLDFFPALYYVLEGEENQDCEFL  
NORSIMEVKRDTDLRLGMRFLGLAKNDIKLTLEVLFEYPVLPYEQKLOHWSLI  
AOVNGSPMDTLDFAFYCLFESODEEYVGLKRFQEWLILINQMDLKVSXYCLKIC  
ONKAIKIVDILDLISVNTLELCPVYTVFQESCKPLIEMWNGNCSYLSIRLKEID  
LGDSILSORAKIICLELRNOSCHTQULTRFSAVYVGLKMLKLSNQKLTNLIG  
NTPMKDDMKLACALHPKCSVETLLDSCELTIGYEWISTLLISTRLKLSLAK  
NRVGVKMSISGNALSSMCLQKLIIDNCGILPASCHLIVSALFSNQNLTHLCLSNP  
SLTEGVQOLCOFLRNEPCALORLIINHCVLDDAYGFLMARLANTKLTHLSLTNP  
VGDGAMLLCEALKEPTCYLOEELVNDCOLTQNCEDLACMITTKQLKSLDGNML  
GDKGVITLCEGLKQNNSSLRRLGKACKLTSNCEALSLAISCPHNSINTLVKNDPS  
TSGMLKICSAFQCEVSNLGIITLWKQRYAVRQLEVEFVKPHVVIDGWTASDED  
DRNWKX"

## ORIGIN

Query Match 31.9%; Score 1287.8; DB 10; Length 3405;  
Beet local Similarity 65.4%; Pred. No. 3.1e-307;  
Matches 1985; Conservative 0; Mismatches 1032; Indels 18; Gaps 6;

Qy 287 AAAAAATTTCAAGCTATGAGAACAGAGGTGCCACAGCAGCAGACAGACAGAACAGAC 346  
Db 256 AAAATGTAGTAAAGCCATCTGTAAGAGCAGTGTATCAGAGAGAGTGAACAGCATCAG 315  
Qy 347 AAATTTCAAGCTATGAGAACAGAGGTGCCACAGCAGCAGACAGACAGAACAGAC 406  
Db 316 AAAAGAAAATGACTTCTCCAGAAAAACGACATTAATCAATCAAAAGAACAGACACAG 375  
Qy 407 ATGAGGTGACACATGGGACTACAAAGTCAAGTATGACCAATTCGCTGAGGAGAG 466  
Db 376 AGCAGAGACACATCAGATTAATGAGAGTACTTAACAAGCTACAAAGCCATGTGATG 435  
Qy 467 ATGTACGTGTGATTTGAAAAACATGCTGTGCTGACGTGCGCGGAATGCAAACTGTGCTG 526  
Db 436 CTAAGTTGACACAAAGTGTGATCTACATGACAGCCAGACATGAATTAATTATGTCTG 495  
Qy 527 GTGCTTTGATTCAGACCGGTGGGGCTTCGGGCTCGCACGCTGCTGTGACGAGAACT 586  
Db 496 ATGCTTTTAAACATACAGAAAACTTCCAGCTCTCACCATTAATCTTACATGGAAGAC 555  
Qy 587 CAGGAATTTGGGAATTCGCTTACCAAGAGATGATGCTGTGGGCGCAAGTGGAC 646  
Db 556 CAGAGTTGGGAAGTCAAGCTTGGCCAGAAATATTTGTTGGGTCAGGCAAGGTAAC 615  
Qy 647 TCTACAGGAGATGTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGACGCGAAGA 706  
Db 616 TCTTCCAAAAATG---TCTTTGTCAATCTTCTTCTGTGAGAAATAAGTGGACAG 672  
Qy 707 AGGAGAGCATGTCAACAGAGTTCACTTCAGAGAGTGGCCAGACTCCAGGCTCCGCTGA 766  
Db 673 AGAAGAGCATGTTGGCAGCTGATTTCTAAGAGAGTCCAGACTCCGAGGATCAGTGA 732  
Qy 767 CGGAGATCATGCTCCGACCGAAGAGGCTGTTCATCATTTGAAGGTTTGGATGACCTGG 826  
Db 733 CAAAGATCATGTCCCAACAGAAAGACTCTGTGTTGTCATAGATGAGCTTGAATATATGG 792  
Qy 827 GCTCTGTCTC---AAGATGACCAAAAGCTCTGCAAAAGCTGAGGAGAGAGCTCTC 883  
Db 793 ACTGTCTCTCAACATGATGATATGACATATCAAGAGACTGGAAGATGAACAGCCCA 852  
Qy 884 CGTTCACTCTCATACGAGTGTGTGAGAGAGAGTCTGCTCTGATGCTCTTCTGATCG 943  
Db 853 TATCATCTGATGATACAGCTCTGAGAGAGGCTTCTCTACCTACGCTCTTCTCTCA 912



OY	944	TCACCGTCAGAGACGTGGGCAACAGAAAGCTCAAGTCAAGAGTGTGTCCTCCGTTAC	1003
Db	913	TTACCACCAAGAAACACAGGCTTAGAAAACTCAAGTCAATGAGTGTTCCTCCCTCTATA	972
OY	1004	TGTTAGTTAGAGGAACTCTCCGGGGAACAAAGATCACTTGCTCTTGAGCGGGGATTG	1063
Db	973	TACTGTTGAAGGACTGTCTGTGATCAAGAGATCTCACTGTGTCTCGAAGAAATCTCA	1032
OY	1064	GTGAGCATCAGAAAGACACAGGCTTGCGTGCATATGACACACCGTAGTCTCGACC	1123
Db	1033	ATGATCTGATGATGAATATACAGCTTCCATTTCTGATAGAAATCAACAGGCTGTTGACC	1092
OY	1124	AGTGCAGGTGCGCGCGGGGCTCTCATCTGCGTGGCCCTGACGCTGACGAGCGTG	1183
Db	1093	AATGCGAGGCCCTCTGTGTGTCTCCCTGTGTGTGAGGCTCTCGAGCTTACAGAAAGAAC	1152
OY	1184	TGGGGAGAGCGTGCGCCCTTTCACCAACACGCTCACAGGCTGACCGCGCTTGTGT	1243
Db	1153	TGGGAAAGATGACCCCTACCTTCGACAGCTCTCACGGTGTGTAATGCCAGTTGGTGT	1212
OY	1244	TTCAATCACTCAACCCCTGAGGCGGTGTCCGGCGCTGTCTCAATCTGAGAGAAAGATG	1303
Db	1213	TTCAACACCTCACTTGAAGAGGCTTCTCCAGAGGCTCTCACTGACGAGAAACAGATTA	1272
OY	1304	TCCTGAACCGCTTCTGCGCGTATGCTGTGAGAGGAGTGGAAATGGAATCAAGTTTG	1363
Db	1273	CTCTAGTGGTTTGTGCATGATGCGAGCTGAAGAGTGTGACCATAGAGTCCGTTCT	1332
OY	1364	ATGTGACGACTCATGCTTCAAGGACTCGGGAGCTGAGCTCCGTCTCTGTTTACA	1423
Db	1333	ATGAATGATGATCTGAAGAACTATAGCTTAAAGGAGTCTGAGATCTTGCCCTCTTACA	1392
OY	1424	TGAACATCCTCTCTCCACAGACGCC--ACTGAGGAGATATACACCTTCTCCACCTCA	1480
Db	1393	TGAACATCTCTCTCAGGTTGGCCACACAGTAGACATGTTATGTTTCTCCCACTCA	1452
OY	1481	GTCTTCAGGACTTCTGTGCGCGCTTGTATCTAGCTTTAGAGGCGCTGGAATCGAGCAG	1540
Db	1453	GCTTCAGGATTTCTTGTCTGCTTATATATATGTTTGAAGGGCTGGA--GGATGGA	1509
OY	1541	CTCTTCGCCCTCTGTACGTTGAGAAAGACAAAGAGTCCATGAGCTTAAACAGCAGCT	1600
Db	1510	ATCAGCATTTTGTCTTCAATTGAAAAACAAAGAGCATATGAGAGTGAAGAACTAGCG	1569
OY	1601	TCCATATCCACTCGCTTTGAGTGAAGAGCTTTCTTGTGGGCTCGTGAGCAGACGTAA	1660
Db	1570	ACACTCGCTC---CTTGGGATGAAGCTTTCTTATTTGGCTCTATGAACAGGATATCT	1626
OY	1661	GGAGGCACTGAGGTCCTGTGCGGCTGTCCGTTCCCTGGGGGTGAAGCAAGCTTC	1720
Db	1637	TGAAGACTCTGGAGGTTCTGTTTGAATATCCCGTATTCCAACTGTTAGCAGAGCTCC	1686
OY	1721	TGCATGAGGTCTCTGTGTGGGTGACGAGCTTATGCCAACACCCAGAGACACCCCTG	1780
Db	1687	AACACTGGGTCTCTGTAGTACGACAGGTCAATGAGCACAAGCCCAATGACACCCCTG	1746
OY	1781	AAGCTTCCACTGTCTTTTCGAGACTCAAGACAAAGATTTGTTCCGCTTGCCATTTAAAC	1840
Db	1747	ATGCTTCTATATGTTATTTAGTCTCAGAGTGAAGAGTTTGTGGCGGGCTCTCAAC	1806
OY	1841	GCTTTCAGAAAGTGTGCTTCGATTTAACAGAACTGGAATTGATAGCATTTCTTCT	1900
Db	1807	GCTTTCAGAAAGTGTGCTGCTGATTTAACAGAAAGTGAATGAGTCTCTTCTTAAT	1866
OY	1901	GCTTTCAGACACTGTCCGTATTTGCGAAATATCCGCTGAGTCAAGAGATTTTCCAA	1960
Db	1867	GCTTCAACACTGTGAGAACTTGAAGGAAATCCGGTGTGATATCAGAGAACTCTCTCG	1926
OY	1961	GAGATAGTCCGCTBAGGACTGTCTGTGTCTCT--CTATGATGCGGATTAAGACC	2017
Db	1927	TAGATTAATCTCTCAGGCTGTGCTCTGTTGTTATCTGTCAAGAGACACAAATGTAAGCCC	1986

QY	2018	TCATTGAGGACACATGGGGAABATTTTCGCTCCATGCTTGGGACCACCCACACCTCGGGC	2077
Db	1987	TCCCTCATTGAGATGGTGGGGAACTTCCTGCTCTGTGCTTGGCGAGCCTCCGGAACCTTGAAG	2046
QY	2078	AGCTGGACCTGAGGAGCAGCATCTCTGACAGAGCGGGCCATGAAGACCTGTGTGCAGAC	2137
Db	2047	AGCTTGACTTGGGGGACACGACATCTTGAATCAACGGGCCATTAAGATACTGTGTGCTCGAGC	2106
QY	2138	TGAGGCATCCCACTGCAAGATACAGACCCCTGATGTTTGAATATGCAAGATTACCCCTG	2197
Db	2107	TGCGGAATCAATCCCTCGACGAATATACGAAGCTGAGGTTTAAAGATGAGAGGATGATGTCTG	2166
QY	2198	GTGTGCAGCACCTCTGAGGAATCGTCATGGCCAAACCTTAACCTTAAGATCCCTCAACTTGG	2257
Db	2167	GCCGTGAACAATCTCTGMAAGCTCCCTTTTATGCAATCAAAACCTTAAAGTAACTCAATCTAG	2226
QY	2258	GAGGACCCCACTGAAGGAAAGAGATGTAAGATGGGGTGTGAAGCCTTAAACACCCCA	2317
Db	2227	GGAACTCTCCATGAAGGATGATGATGAAATTAGCTGTGCAAGCCGTGAACATCTCCA	2286
QY	2318	AATGTTGTTGGAGTCTTTTGAAGCTGATTTGCTGTGATTTGACCAATGCTCTGTACTCTGA	2377
Db	2287	AGTGTCCGTGTGAGACTCTGAGTTTGATTCCTGTGATTTAACATATATGGCTATGAGA	2346
QY	2378	AGATCTCCCAATCTCTTACGACTTCGCCACGCTGMAATCTGTAGCTGTGCAGAAACA	2437
Db	2347	TGATCTCCACGCTTCTTATTTTCAACCAACAGGCTAAAGTGTCTGACGCTGGCCAAAATA	2406
QY	2438	AGGTGACGACCAAGGAGTATATGCTCTCAATGATATGCTTGAAGAGTCTCCCAAGGCGCC	2497
Db	2407	GAGTGGAGTAAAGACATGATATCCCTTGGAAATGCTTGAATGACTCAATGTGTCTAC	2466
QY	2498	TGCAGAAAGCTATATCTGAGAGACTGTGGCATCACAGCCACGGGTTTGCAGAGTCTGCGCT	2557
Db	2467	TGCAAAAGTTATATCTGGACAACTGTGGCTTCAACCTGCAGACTGCCAACCCTTGTGCT	2526
QY	2558	CAGCCCTCTGCAGCAACCGAGCTTTGACACACTGTGCTATTCACAACAACGCTTGGGGA	2617
Db	2527	CAGCCCTTTTACAGCAACAGAACTTGGACACACTGTGCTGTCAAAACAACGCTTGGGGA	2586
QY	2618	ACGAAGGTATTAATCTATCTGTGTGATTCATATGAAGCTTCCCACTGTATGTGTGAGAGGC	2677
Db	2587	CTGAAGGATGTCACAAGCTGTGTGATTCCTGAGAAATCCGAATGTGCTCTCCAGCGGC	2646
QY	2678	TGATGCTGAATCAGTGCACCTGTGACAACGCGCTGGCTGTGTTTCTTGCACTTGCCTTA	2737
Db	2647	TGATATCTGAATCATCTGCAACATTTGTATGATATGCTTATGTGCTTCTGTGCAATGAGACTTG	2706
QY	2738	TGGGTAACTCATGCTGACGCACTGACCTTACATGAACCTGTGTGAAGACAAATGGCG	2797
Db	2707	CAAAACAACAAGACTGACCCACCTGAGCCTGACATGAACCCGCTAGGGGATGTGTGCA	2766
QY	2798	TGAAGCTTCTGTGGAAGTCAATGAGAAACCATTTGTTCATCTGCAGACCTGGAATTTGG	2857
Db	2767	TGAAGCTACTGTGTGAAGCTTTAAAGAACCTTACTTGTACTTCAAGAACTGGAACCTAG	2826
QY	2858	TAAAGTGTCACTACACGCGCGCTGTGTGAGATCTGTCCCTGTGTATCTGCAGAGACA	2917
Db	2827	TGACATGCGCACTACACAGAACTGTGTGCGAGACCTGGCCTGTATGATTCACAACAACA	2886
QY	2918	GACACCTGAAGAGCCTGATCTCAACGACAAATGCCCTGTGTGACGATGGGGTTTGTGCGC	2977
Db	2887	AGCAATTAATAAAGTTTGTGATCTTGTGTAAACAACGCTTGCGGTGAACAAGGATCAATTAACC	2946
QY	2978	TGTGTGAGGGACTGAAGCAAAAGAAACAGTGTCTTGACAGACCTCGGGTTGAAGCATGTG	3037
Db	2947	TGTGTGAGGGACTGAAGCAAAATTAACAGCTTCCCTGAGGAAGCTTGGGTGTGGGGCATGTA	3006
QY	3038	GACGTACCTTCGATTTGTGTGAGGGAACCTTTCCTTGGCCCTTTCCTGCAACCGGATCTGA	3097
Db	3007	AGTTGACTTCAATTTGCTGTGAGGCAATGTTCATTTGGCAATCTCTTGTGCAACCTCTCACTGA	3066
QY	3098	CCAGTCTAAACCTGTGTGACAAATTAATCTCAATGCCCAAGAAATGATGAAGCTGTGTTCCG	3157

	FEATURES	source
Dn	3067 ACAGCTTAACCTGGTGAAGAATGACTTCATGACATCGGGAGATGTGAAGCTTGTCCTG	3126
Oy	3158 CTTTGCCTGTCGCCACGTCCTAACCTTACAGATAATTGGGCTGTGGAAATGGCAGTACCCTG	3217
Dn	3127 GATTCCATGAGCCCTGTCTCTCAACCTGGGAGATTAATTGGCTGTGGAAAGCAGGATCATATG	3186
Oy	3218 TGCATAATAGGAAGTCTGTGGAGAAATGCGACGCTACTAACCCTCCGAGTGGTAATTGACG	3277
Dn	3187 CCCGAGTAGAAGAACAAGCTGGAGGAAAGTGAAGTTGTCMAAGCCCACTGGTGAATTGATG	3246
Oy	3278 GTAGTGGCATTTCTTTTGTATGAAGATGACCCGAC 3312	
Dn	3247 GTGATGGTATGCTAGTATGATGAATGACCCGAAAC 3281	
RESULT_15	AY329484 3432 bp mRNA linear ROD 18-DEC-2003	
LOCUS	Mus musculus strain SMR/J MATRR protein isoform-E (Mater) mRNA,	
DEFINITION	complete cds; alternatively spliced.	
ACCESSION	AY329484	
VERSION	AY329484.1 GI:37595466	
KEYWORDS	.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3432) Cheng, H., Zhang, Y., Huang, Z. and Ma, R.Z. Identification and characterization of alternative splice variants for murine Mater Unpublished 2 (bases 1 to 3432) Cheng, H., Huang, Z., Zhang, Y., Liu, H., Teuscher, C. and Ma, R.Z. Direct Submission Submitted (20-JUN-2003) Human and Animal Genetics Division, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Building 917, Datun Road, Beijing 100101, P.R.China location/Qualifiers 1..3432 /organism="Mus musculus" /mol_type="mRNA" /strain="SMR/J" /db_xref="taxon:10090" 1..3432 /gene="Mater" 9..3296 /gene="Water" /note="alternatively spliced" /codon_start=1 /product="water protein isoform-E" /protein_id="AAO94607.1" /db_xref="GI:37595467"	
gene	/translation="MGPEPKSKAILKARGIEBQKSESTSPENSRALLDQSGEEVEQASQEQFSESTMGPPEKDSKAILKARGIEBQKSESTSPENSRALLDQSGEEVEQASERKMTSPEDSKSIQDKQGEQEBQETSDNGDLQYKAHIAKPTSYDLHYDEPMKL LSAFKPYOKTFQPHITILIHGRPGVGSALARSIVLGMAGKGLFOKMSFYLFSEIREIKMTSSSLAOLIAKECPDSMDLVTKINSOPERRLFEVIDGDMDSVLOHDMTSIRMPKDQPIYIAMSILRKALPPSPLIITTRNTGELKASMVSPLIYIVBGLSARBOQLVLENIEMESRIQVFSLIHNDLPDQCQAPSVCSLVCEALQJQDKLGKRTLPOTQTLKEEIALFMNILLQVGHNSBOCYASHLSLDYFAVALYYEGLEBNMOHCFLIEINORISMEYRTDDTRLIGMKRFPLGNANKDILKTLEVLFEYPVLPYEOKLQHWVSLILEAOQNYESPMDTDAFYCLPESODEEPFGALKLRPOEVMLINOKMLKYSSVCYLKC ONLKAIRVDIDILSDVTIELCYVVYVOEFQCKPLMBWGNFCSVGISRLNKELD LGDISLSQRAMKILCLELRNQSCRIOKLTESAASVSGAKHLMWLFSNONUKYLNKG NTIPMKDDMKLACALKHKPCQSVETLRISCETLIIGEMISTLLISTTRKCLSLKAN NRVAQNMISLGNALSSMCLOKLLIDNCGLTASCHLVSALFSNONITHLCLSNKN SLTEGVGOOLCOPLRNPECALOILIHNCHNVDAAYGTLAWRLANNKLTHTLSLTMNP VGDGAMTLCALKEPCTYOELVELVDOULQONCEDDLACITTYTHKLKSIDLGNAL AGDKATITLCBGKQSSSLRLRGJAGCALISNCEBALISLCPHILNSILVNDSF TSQMLKICSAQCEPVSNLGIILGWKOEYAVRAVRKEVEKVPVHILDGWVASDBD	
CDS		

ORIGIN	DRNMWKN"
Query Match	31.9%; Score 1286.2; DB 10; Length 3432;
Best Local Similarity	65.4%; Pred. No. 7.7e-307;
Matches 1984; Conservative	0; Mismatches 1033; Indels 18; Gaps 6;
QY	AAAAATTTCACAACTATGGAACAAGAAAGTGGCCACACGACGACAGACAGAAACAG 346
DB	287 AAAAAATTTCACAACTATGGAACAAGAAAGTGGCCACACGACGACAGACAGAAACAG 346
QY	256 AAAATGTCACTAGAGCAATCTGAAGAAGACGTGATCGAAGAAATGGAACGCGTCAG 315
DB	347 AAAATTCACAAAGTATGGAACAAGAAAGTGGCCACGACGACAGACAGAAACAGAGAC 406
QY	316 AAAAGAAAAAGACTTCTCCAGAAAACGACAGTAATCAATCCAGAAAAGACCAAGACAG 375
DB	407 ATGAGAGTGCACATGAGGACCTACAAAGAGTCAGGTATGACCAAAATTCGCTGAGAGAGAG 466
QY	376 AGCAGAGACGACATCAGATTAATGAGAGTGACTTACAAAGACTACAGGCCCATGTGATTG 435
DB	467 ATGTAAGTGTGATTTGAAAAACCTGCTGCTGATCGTCGCGGAATGCAACGTTGGCTG 526
QY	436 CTAAAGTTCGACACAAAGTGTGATCTACACTAAGACAGCCACAGAGTAATATTGTCCTG 495
DB	527 GTGCTTTTGAATTCAGACCGGTGGGGCTTCGGGCTCGACCGGTGGTTCGACAGGAAAGT 586
QY	496 ATGCTTTTAAACATACACAAAACTTTCAGCTTCACACCAATTATCTTAATGGAAGAC 555
DB	587 CAGGAATTTGGAATATGCGCTTACGACAGAAAGATGTCGTGCTGGCGCAAGGTGAC 646
QY	556 CAGGAGTTGGAGATCAGCTTTGGCCAGAAAGTAATGTTCTTGGCTGGGACAGGGTAAAC 615
DB	647 TCTACAGGGAATGTTCTCTTACGTCCTTCTCTCCCGTTAGAGATGCAAGCGGAAGA 706
QY	616 TCTTCCAAAAAATG--TCTTTTGTATCTTCTCTGTTAGAAAAATTAAGTGGACAG 672
DB	707 AGGAGAGAGGTGCACAGAGTTCACTCCAGGGAATGGACAGATCCCAAGGCTCCGGTGA 766
QY	673 AGAAGAGAGTTGGACAGCTGATTTGCTTAAGAAAGTGCAGACATCTGGGAATCTAATGA 732
DB	767 CGGAGATCATATGTCCTCCAGACCAAGAAAGCTGTTGTTCAATTTGAAGACCTGG 826
QY	733 CAAAGATCATGTCCTCCAAACCAAGAAAGCTTTGTTGTAATGATGGCTTGATGATAG 792
DB	827 GCTCTGCTCTC--AACAATGACACAAAGCTTCGAAAAGCTGGCTGAGAAAGCAGCTTC 883
QY	793 ACTCTGTCTCCCAACATGATGATGATGACACTATCCAGAGACGTGAAAGATGAACAGCCCA 852
DB	884 CGTTCAACCCATACAGCAATCTGCTGAGGAAGAGTCTCTGCTCCGATGTCCTTCCGATCG 943
QY	853 TATACATCTCGAATGTAACGCTCTCTGAGGAAGGCTCTCTCACTCAAGTCTTTTCTCATCA 912
DB	944 TCACCGTCAGAGACGTGGGACACAGAGAGCTCAAGTCAAGAGTCTGTCCTCCGTTAAC 1003
QY	913 TTTACACACGAAACACAGGCTTTAGAAAACTCAATCATGGTTGTGTCCCTCTATATA 972
DB	1004 TGTTAATTAGGAAGATCTCCGGGGAACAAAGATCCACTTGTCTCTTGAAGCGCGGAGATTG 1063
QY	973 TACTGTTTGAAGGACGTCTGATCAAGAGAGATCTCAGCTGGTCTCTGAGAAACATCTCCA 1032
DB	1064 GTTAGCATCAGAAAGACAAAGGTTGCGTGCAATATGAACAAACCGTGAAGCTGTCTCCACC 1123
QY	1033 ATGAGTCTGATTAATAATCAAAAGTCTTCATCTCTATAGAAANTCAACCAAGCTGTTTGACC 1092
DB	1124 AGTGCAGAGTGCCCGCGCGTGGGCTCTCATCTGGGTGAGCCCTGACGCTGACAGACGTTG 1183
QY	1093 AATGTCAGGCGCCCTCTGTGTGTCTCTGTCGTGTGTGAGAGCTTCGACGCTTAACAGAAAGAAC 1152
DB	1184 TGGGAGAGAGCGCTCGCCCTTTCAACCAACGCTTCACAGGCTGACCGCCGCTTTTGTGT 1243
QY	1153 TGGGAAAGAGATGACACCTTAACCTTGACAGATCTCACCGGTTTGTATGACCAAGTTGGTGT 1212
DB	1244 TTCAATCAGCTCACCCCTCGAGGCGTGGTGTCCGCGCTGTCTCAATCTTGGAGAAAGAGTTG 1303

Db 1213 TTCACAGCTCACTTGAAGAGGCTTCCAGAGCGCTCTCAGTCAGAGAAAGACATTA 1272  
 Qy 1304 TCCGAGCGCTTTCGCCCTATGCTGTGGAGAGAGTGTGAATAGAAAGTACAGTTTG 1363  
 Db 1273 CTCTAGTGGTGTGATGATGAGGAGCTGAAGAGTGTGACATGAGTGGTGTCT 1332  
 Qy 1364 ATGTGAGAGCTCAATGTTCAAGAGCTGGGAGTGTGAGCTCGGTCTGTTTACA 1423  
 Db 1333 ATGATGATGATGAAAGACTATAGCTTAAAGAGTGTGAGTCTTGCCCTCTTACA 1392  
 Qy 1424 TGAACATCTCTTCCCAAGACCC--ACTGTGAGAGTACTACACTTCTTCCACTCA 1480  
 Db 1393 TGAACATCTCTTCCAGGTTGGCCCAACAGTGAAGAGTGTATGTTTCTCCCACTCA 1452  
 Qy 1481 GTCTCAGAGACTTGTGTGGGCTGTACTAGTGTGAAGGCTGTGAATTCAGCCAG 1540  
 Db 1453 GCTTCAGAGATTTCTGTGCTTATATATATATATATATATATATATATATATATAT 1509  
 Qy 1541 CTCTCTGCCCTCTGATCTGATGAGAGACAAAGAGTCCATGAGCTTAAACAGGAGCT 1600  
 Db 1510 ATCAGACTTTTGTCTTCAATGAAACCAAGAGCATCTGAGGTGAAGAGAACTGAG 1569  
 Qy 1601 TCCATATCCACTCGCTTGTGATGAAGCGTTTCTTGTGGCCTGTGAGCGAAGCTGA 1660  
 Db 1570 ACACTGCGCTC--CTTGGATGAAGCGTTTCTTATTTGGCTCATGAACAGATATCT 1626  
 Qy 1661 GAGGCGCACTGAGAGTCTGTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGCTTC 1720  
 Db 1627 TGAACATCTGTGAGGTCTGTGTGAATATCCGTATTCMACTGTGAGCAAGAGCTCC 1686  
 Qy 1721 TGCACTGGGCTCTCTGTTGGTGTGAGCGAGCTATAGCCACCAAGGAGAGACCGTGG 1780  
 Db 1687 AACATCTGGTCTCTGTATAGCTCAGAGGTCAATGAGTCAAGCCCAATGAGACCGTGG 1746  
 Qy 1781 AGCGCTTCACTGTCTTTCGAGCTCAAGACAAAGATTTGTGCTGTGAGCTTAAACA 1840  
 Db 1747 ATGCTTGTATGTCTATTTGAGTCTCAGAGTGAAGATTTGTGGGGGCTCTCAAC 1806  
 Qy 1841 GCTTCAAGAGTGTGGCTTCCGATTAACAGAACCTGATCTTGAATCTTCTTCT 1900  
 Db 1807 GCTTCAAGAGTGTGGCTGTGATTAACAGAACATGATCTTGAAGGCTCTCTCACT 1866  
 Qy 1901 GCTTCAGAGCTGTGCTATTTGGGAAATTCGGGTGATGTCAAGGAGCTTCCCA 1960  
 Db 1867 GTCTTAAGACTGTCAAGCTTGAAGCAATCCGGTGAATACAGAGCTCTCTCTGG 1926  
 Qy 1961 GAGATGAGTCCGCTGAGGATGTCTGTGTCTCT--CTATGATGTGGGATTAAGACC 2017  
 Db 1927 TAGATATATCTCTGAGGCTGTGCTGTGTGTACTGTCCAGAGACAAATGTAAGCCC 1986  
 Qy 2018 TCATTTGAGAGAGTGTGGAAATTTCTGTCTCAATGCTTGGACCAACCACTGTGGG 2077  
 Db 1987 TCCCTATGAGTGTGGGAAATCTTGTGTCTGTCTGTGGCAGCTCCGGAATCTGAAG 2046  
 Qy 2078 AGCTGAGCTGTGGGAGAGACTCTGACAGAGAGGGGCAATGAAGCCGTGTGGCAAGC 2137  
 Db 2047 AGCTGAGCTGTGGGAGAGACTCTGAGTCAACGGGCAATGTAAGTCTGTGGC 2106  
 Qy 2138 TGAGGATCTCCACTGCAAGATACAGACCTGATTTTGAATCAACAGATTAACCCCTG 2197  
 Db 2107 TGGGATATCAGTCTGTGAGAAATACAGAGCTTAAAGTGTGAGAGTGTGTCTG 2166  
 Qy 2198 GTGTGAGAGCTGTGTGAATGTCTATGAGCCCAACCTGAACCTTAAGATCCCTCACTGG 2257  
 Db 2167 GCTTGAACATCTGTGAAGCTCTTTTATGCAATCAAACTTAAGTCACTCAATGAG 2226  
 Qy 2258 GAGGACCACTCTGAAGAGAGAGATGAAGATGGCGTGAAGCTTAAACCAACCCCA 2317  
 Db 2227 GGAACATCTCTGAAGAGATGAAGATGAAGTGAAGCTGTGCAAGCGCTGAACCTCA 2286  
 Qy 2318 AATGTTTGTGAGTCTTGAAGCTGATTTGCTGTGATTTGACCAATGCTGTACTGA 2377  
 Db 2287 AGTGTCTGTGAGACTGTGAGTGTGATTTCTGTGATTAACATATGTTATGAGA 2346

Qy 2378 AGATCTCCAAATCTCTTGAACCTTCCGAGGCTGAATATCTGTAGCCGTGGAGAAACA 2437  
 Db 2347 TGAATCTCAAGCTTCTTATTTTCAACACAGGCTTAAAGTGTCTGAGCTGTGCAAAATA 2406  
 Qy 2438 AGGTGACAGACGAGGAGATATGCTCTCAGTATGAGTGTGAGAGTCTCCAGTGGGCC 2497  
 Db 2407 GAGTGGAGATTAAGATGATATATCTTGGGAATGCCCTGTGATGATCTCAATATGTCTAC 2466  
 Qy 2498 TGCAGAGCTGATATCTGAGAGACTGTGATCAACAGCAAGGAGTGTCCAGAGTCTGGCT 2557  
 Db 2467 TGCAGAGTGTATCTGAGACACTGTGCTCAGCACTGTGAGCTGCAACCTTCTGTGCT 2526  
 Qy 2558 CAGGCTGTGAGCAACGGAGCTTGAACAACCTGTGCTATTCAGAACAGGCTGGGGA 2617  
 Db 2527 CAGGCTTTTTCAGCAACAGAACTTGAACAACCTGTGCTGTCAACAGGCTGGGGA 2586  
 Qy 2618 ACGAAGTGTAAATCTACTGTGTGATCAATGAGGCTTCCCACTGTATGTGAGAGGCT 2677  
 Db 2587 CTGAAGAGTGTGAACAGCTGTGTGATCTGTGAGGAAATGAGATGTCTCCAGCGGC 2646  
 Qy 2678 TGAATCTGAATCAGTGCACCTGAGACAGGCTGTGTGTGTTTCTTGAACCTTGGCTTA 2737  
 Db 2647 TGAATCTGAATCAGTGCACCTGAGACAGGCTGTGTGTGTTTCTTGAACCTTGGCTTA 2706  
 Qy 2738 TGGTAACTCAATGCTGAGAGAGCTGAGGCTTGAATGAATGAATGAGCAATGGCG 2797  
 Db 2707 CAACAGACAAAGTGTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGAGTGTGAA 2766  
 Qy 2798 TGAAGCTTGTGAGGAGTCAATGAGAGAACTTGTGATCTTCCAGAGCTGTGAGTTGG 2857  
 Db 2767 TGAAGCTTGTGAGGAGTCAATGAGAGAACTTGTGATCTTCCAGAGCTGTGAGTTGG 2826  
 Qy 2858 TAAAGTGTATCTCAACGCGGCTGTGTGAGAGTGTGTCTGTGTGATCTTCCAGAGCA 2917  
 Db 2827 TGAAGCTTGTGAGGAGTCAATGAGAGAACTTGTGATCTTCCAGAGCTGTGAGTTGG 2886  
 Qy 2918 GACACTGAAGAGCTGTGATCTCAAGAGCAATGCTGTGTGTGAGAGTGTGTGTGAG 2977  
 Db 2887 AGCACTTGAAGAGTGTGATCTTGTGTGAACAGGCTGTGTGTGAAGAGTGTGTGA 2946  
 Qy 2978 TGTGAGAGGAGCTGAAGAGAAAGAAAGTGTGTGAGAGAGCTGTGTGTGAGAGGATGTG 3037  
 Db 2947 TGTGAGAGGAGCTGAAGAGAAAGTGTGTGAGAGAGCTGTGTGTGAGAGGATGTG 3006  
 Qy 3038 GACTGACTTGTGATGTGTGAGAGGACTGTCTGTGTGAGGCTTGTGTGAGAGGATGTG 3097  
 Db 3007 AGTGAATCTTCAATGTGTGAGGAGTGTGTGTGAGGAGCTGTGTGAGAGGATGTG 3066  
 Qy 3098 CCAATCTGAAGCTGTGAGAGAAATCTTCAATGAGGAGGATGTGTGTGAGAGGATGTG 3157  
 Db 3067 ACGAGCTGAAGCTGTGAGAGAAATCTTCAATGAGGAGGATGTGTGTGAGAGGATGTG 3126  
 Qy 3158 CTTTGTGAGGAGCTGTGAGAGAAATCTTCAATGAGGAGGATGTGTGTGAGAGGATGTG 3217  
 Db 3127 GTTTCATATGCTGTGTGTGAGAGAAATCTTCAATGAGGAGGATGTGTGTGAGAGGATGTG 3186  
 Qy 3218 TGAATTAAGAGAGTGTGAGAGAAATCTTCAATGAGGAGGATGTGTGTGAGAGGATGTG 3277  
 Db 3187 CCGAGATGAAGAGAGTGTGAGAGAAATCTTCAATGAGGAGGATGTGTGTGAGAGGATGTG 3246  
 Qy 3278 GTAGTGTGATCTTGTGTGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3312  
 Db 3247 GTAGTGTGATCTTGTGTGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3281

Search completed: July 18, 2005, 22:13:41  
 Job time : 17042 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 12:25:31 ; Search time 2009 Seconds  
(without alignments)  
11889.575 Million cell updates/sec

Title: US-10-066-521-5  
Perfect score: 4035

Sequence: 1 atggaagagagcaaatcgct.....tcgaggggctgtgtctctaa 4035

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_1Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4035	100.0	4035	6	AA144356 Human PYR
2	4035	100.0	4035	10	ADP94744
3	3166.2	78.5	3900	6	ABK48628 Human MAT
4	3166.2	78.5	3900	6	AD49018 Human MAT
5	3139	77.8	3830	8	ADA45220 Human MAT
6	3133.6	77.7	3489	6	AA41224 Human EMB
7	3133.6	77.7	3926	8	ADA45218 Human MAT
8	3057.8	75.8	5859	6	AA147135 Human MAT
9	3057.8	75.8	6939	6	AA147131 Human MAT
10	3057.8	75.8	6939	6	AA147131 Human MAT
11	2900.6	71.9	3226	6	ABK48610 Human MAT
12	1840.6	45.6	2820	12	ADN62078 Human CDN
13	1278	31.7	3447	8	ABK48611 Human MAT
14	1278	31.7	3447	8	AD49001 Human MAT
15	1152.8	28.6	1157	8	ABK48609 Human MAT
16	1152.8	28.6	1157	8	AD48999 Human MAT
17	756.6	18.8	1075	8	ABK48610 Human MAT
18	756.6	18.8	1075	8	AD49000 Human MAT
19	536	13.3	2808	12	ADP47735 Human inf
20	527.8	13.1	3039	12	ADP47733 Human inf

21	440	10.9	2775	12	ADO50889 Human NOV
22	438	10.9	2775	10	ADP44704 Human NOV
23	424	10.5	2985	6	ADP36408 Human PAA
24	424	10.5	2985	12	ADJ19330 Human PAN
25	422.4	10.5	2982	10	ADP94764 Human PYR
26	422.4	10.5	3190	10	ADA53511 Human cod
27	422.4	10.5	3368	6	AA144366 Human PYR
28	422.4	10.5	3368	10	ADP94760 Human PYR
29	418.2	10.4	2575	6	AA170683 Human nuc
30	418.2	10.4	2767	8	ABT16018 Human relA
31	419.2	10.4	3172	6	ABN99366 Human sec
32	419.2	10.4	3310	13	ACN40690 Human sec
33	418.8	10.4	2883	12	ADO41637 Novel hum
34	409	10.1	3186	6	AA144363 Human PYR
35	409	10.1	3186	10	ADP94754 Human PYR
36	404.4	10.0	3108	6	ADP36416 Human PAA
37	404.4	10.0	3108	12	ADJ19338 Human PAN
38	404.4	10.0	3218	12	ADJ19382 Human PAN
39	404.4	10.0	3300	6	AA147129 Human PAN
40	404.4	10.0	3731	12	ADP47723 Human Mon
41	397.2	9.8	2799	12	ADH18983 Human cel
42	394.2	9.8	3531	6	AA147128 Human cel
43	392.6	9.7	3189	5	AA167185 Nucleotid
44	392.6	9.7	4944	10	ADP14464 Human end
45	381.4	9.5	3431	4	AD14322 Human nuc

## ALIGNMENTS

RESULT 1	AA144356	standard; CDNA; 4035 BP.
ID	AA144356	standard; CDNA; 4035 BP.
XX	AA144356	(first entry)
AC	AA144356	
XX	31-OCT-2002	
DT	31-OCT-2002	(first entry)
XX	Human PYRIN-5	cdna sequence.
DE	Human PYRIN-5	cdna sequence.
XX	Human; gene; ss; gene therapy; PYRIN; stress-related response;	
KW	apoptotic response; inflammatory response; inflammatory disorder;	
KW	immune system disorder; Crohn's disease; multiple sclerosis; cancer;	
KW	leukemia; autoimmune disorder; arthritis; neurological disease;	
KW	Alzheimer's disease; Parkinson's disease; chromosomal mapping;	
KW	tissue typing; forensic biology; predictive medicine; pharmacogenomics;	
KW	transcription profiling; PYRIN-5.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	CDS	1..4035
FT		/*tag= a
FT		/product= "Human PYRIN-5"
XX	WO200261049-A2.	
XX	08-AUG-2002.	
XX	31-JAN-2002; 2002WO-US002967.	
XX	31-JAN-2001; 2001US-0265231P.	
XX	10-SEP-2001; 2001US-0318645P.	
XX	(MIL-) MILLENNIUM PHARM INC.	
XX	(AMHP ) WYETH.	
XX	Bertin J, Wang W, Blatcher W,	
XX	WPI; 2002-627477/67.	
XX	P-PSDB; AAO15585.	
XX	New PYRIN polypeptides and nucleic acids useful for modulating and	

PT diagnosing stress-related, apoptotic and inflammatory responses, or for  
PT treating inflammatory and immune system disorders, cancers, or  
PT neurological diseases.

PS Claim 4; Fig 3; 167pp; English.

XX The invention comprises the amino acid and coding sequences of human  
XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are  
XX useful for modulating and diagnosing stress-related, apoptotic and  
XX inflammatory responses. The PYRIN protein and DNA sequences are useful  
XX for treating: inflammatory disorders and immune system disorders (e.g.  
XX Crohn's disease, reactive arthritis, multiple sclerosis, contact  
XX dermatitis, psoriasis, graft rejection, allergies, viral infections and  
XX bacterial infections); cancer (e.g. leukemia); autoimmune disorders  
XX (e.g. systemic lupus erythematosus and arthritis); and neurological  
XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN  
XX protein and DNA sequences may also be used in screening assays, detection  
XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),  
XX predictive medicine (e.g. diagnostic assays, clinical trials and  
XX pharmacogenomics) and transcription profiling. The present DNA sequence  
XX encodes the human PYRIN-5 protein

SO Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;

Query Match 100.0%; Score 4035; DB 6; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAAGAGCAAAATCGCTCAAGCTTTTCAGCTACGGGCTGCAATGCTGCTATAG 60
DB 1 ATGAAAGAGCAAAATCGCTCAAGCTTTTCAGCTACGGGCTGCAATGCTGCTATAG 60
QY 61 CTAGCAAGAGAAATTCAGCATTCAGAAATTAAGAAATTCAGAAATTCAGAAATTC 120
DB 61 CTAGCAAGAGAAATTCAGCATTCAGAAATTAAGAAATTCAGAAATTCAGAAATTC 120
QY 121 ACCACATGCTCTATTCACAGTTGAAATTCAGAAATTCAGAAATTCAGAAATTC 180
DB 121 ACCACATGCTCTATTCACAGTTGAAATTCAGAAATTCAGAAATTCAGAAATTC 180
QY 121 ACCACATGCTCTATTCACAGTTGAAATTCAGAAATTCAGAAATTCAGAAATTC 180
DB 121 ACCACATGCTCTATTCACAGTTGAAATTCAGAAATTCAGAAATTCAGAAATTC 180
QY 181 CTCTTGCATGATATTATGAGCATCGCTGGCTGGCTGAGTCAATGATGATCTTTGAA 240
DB 181 CTCTTGCATGATATTATGAGCATCGCTGGCTGGCTGAGTCAATGATGATCTTTGAA 240
QY 181 CTCTTGCATGATATTATGAGCATCGCTGGCTGGCTGAGTCAATGATGATCTTTGAA 240
DB 181 CTCTTGCATGATATTATGAGCATCGCTGGCTGGCTGAGTCAATGATGATCTTTGAA 240
QY 241 AACATGAACCTGCGAACCCTCTCGAGAAAGGACGGGATGACATGAAAAAATTTCA 300
DB 241 AACATGAACCTGCGAACCCTCTCGAGAAAGGACGGGATGACATGAAAAAATTTCA 300
QY 241 AACATGAACCTGCGAACCCTCTCGAGAAAGGACGGGATGACATGAAAAAATTTCA 300
DB 241 AACATGAACCTGCGAACCCTCTCGAGAAAGGACGGGATGACATGAAAAAATTTCA 300
QY 301 GCTATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 360
DB 301 GCTATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 360
QY 301 GCTATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 360
DB 301 GCTATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 360
QY 361 ATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 420
DB 361 ATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 420
QY 361 ATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 420
DB 361 ATGGAACAAGAGGTGCGACAGCAGACAGACAGAAACAAAGAAATTTCAAGCT 420
QY 421 TGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGGAGGATGATGCTGATG 480
DB 421 TGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGGAGGATGATGCTGATG 480
QY 421 TGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGGAGGATGATGCTGATG 480
DB 421 TGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGGAGGATGATGCTGATG 480
QY 481 TTTGAAAACTGCTGCTGACCTGGCCGGAATTCGAAAGTGGCTGCTTTGATTC 540
DB 481 TTTGAAAACTGCTGCTGACCTGGCCGGAATTCGAAAGTGGCTGCTTTGATTC 540
QY 481 TTTGAAAACTGCTGCTGACCTGGCCGGAATTCGAAAGTGGCTGCTTTGATTC 540
DB 481 TTTGAAAACTGCTGCTGACCTGGCCGGAATTCGAAAGTGGCTGCTTTGATTC 540
QY 541 GACCGGTGGGCTTCGCGCTCGACAGGTGTTCTGACGGAAGTCAAGAAATTTGG 600
DB 541 GACCGGTGGGCTTCGCGCTCGACAGGTGTTCTGACGGAAGTCAAGAAATTTGG 600
QY 541 GACCGGTGGGCTTCGCGCTCGACAGGTGTTCTGACGGAAGTCAAGAAATTTGG 600
DB 541 GACCGGTGGGCTTCGCGCTCGACAGGTGTTCTGACGGAAGTCAAGAAATTTGG 600
QY 601 TCGGCTGAGCAGAAAGTCTGCTGCTGGGCGCAAGGTGACCTTACAGGAAATG 660
DB 601 TCGGCTGAGCAGAAAGTCTGCTGCTGGGCGCAAGGTGACCTTACAGGAAATG 660
QY 601 TCGGCTGAGCAGAAAGTCTGCTGCTGGGCGCAAGGTGACCTTACAGGAAATG 660
DB 601 TCGGCTGAGCAGAAAGTCTGCTGCTGGGCGCAAGGTGACCTTACAGGAAATG 660
QY 661 TTCTCTACGCTCTTCTCTCCGTTAGAGATGACGCGAAGAAAGAGACAGTGT 720
```

```
DB 661 TTCTCTACGCTCTTCTCTCCGTTAGAGATGACGCGAAGAAAGAGACAGTGT 720
QY 721 ACAGATTCATCTCCAGAGATGGCCAGACTCCAGGCTCCGGTACAGAGATCATGTC 780
DB 721 ACAGATTCATCTCCAGAGATGGCCAGACTCCAGGCTCCGGTACAGAGATCATGTC 780
QY 781 CGACCAAGAAAGGCTGTTGTTTCATGATGACGTTTCATGACCTGCTGCTCAAC 840
DB 781 CGACCAAGAAAGGCTGTTGTTTCATGATGACGTTTCATGACCTGCTGCTCAAC 840
QY 841 AATGACAAAGCTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 900
DB 841 AATGACAAAGCTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 900
QY 901 AGTGTGCTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 AGTGTGCTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GGCAACAGAAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1020
DB 961 GGCAACAGAAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1020
QY 1021 TCGGGGAAACAAAGATTCACCTGCTCTTGAAGCGGGATTTGATGACATCAGA 1080
DB 1021 TCGGGGAAACAAAGATTCACCTGCTCTTGAAGCGGGATTTGATGACATCAGA 1080
QY 1081 CAAAGGCTGCTGATCAATGAAACACCGTGAAGTCTGCAACCGTCAAGTCAAG 1140
DB 1081 CAAAGGCTGCTGATCAATGAAACACCGTGAAGTCTGCAACCGTCAAGTCAAG 1140
QY 1141 GTGGGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 GTGGGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 CCTTTCAACAAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1260
DB 1201 CCTTTCAACAAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1260
QY 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 CGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
DB 1321 CGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1381 GTTCAAGGATCTGGGAGTCTGAGCTCGTGTCTGTTTCAATGAAATCTCTTCCA 1440
DB 1381 GTTCAAGGATCTGGGAGTCTGAGCTCGTGTCTGTTTCAATGAAATCTCTTCCA 1440
QY 1441 GACAGCACTGTGAGGATCTGACCTCTTCTTCACTGCTGCTGCTGCTGCTGCT 1500
DB 1441 GACAGCACTGTGAGGATCTGACCTCTTCTTCACTGCTGCTGCTGCTGCTGCT 1500
QY 1501 GCTTTGACTACGCTGTAGAGGCTGGAATTCAGACAGCTCTGCTGCTGCTGCT 1560
DB 1501 GCTTTGACTACGCTGTAGAGGCTGGAATTCAGACAGCTCTGCTGCTGCTGCT 1560
QY 1561 GAGAAACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCATATCCACTGCTTGG 1620
DB 1561 GAGAAACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCATATCCACTGCTTGG 1620
QY 1621 ATGAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 ATGAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1681 CTGGGCTGTCCCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1681 CTGGGCTGTCCCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 1741 GGTGAGAGCTTAAATGCAACCCAGAGACACCTTGAAGCTTCACTGCTTCTTTC 1800
DB 1741 GGTGAGAGCTTAAATGCAACCCAGAGACACCTTGAAGCTTCACTGCTTCTTTC 1800
```



Db 1741 GGTTCAGACGCTTAATGCCACCAACCCAGAGAGACCTCTGAGCGCTTCACTGTCTTTTC 1800  
Qy 1801 GAGACTCAAGAACAAAGAGTTTGTGGCTTGGATTAAACAGTTTCCAGGAATGTGGCTT 1860  
Db 1801 GAGACTCAAGAACAAAGAGTTTGTGGCTTGGATTAAACAGTTTCCAGGAATGTGGCTT 1860  
Qy 1861 CCGATTAAACAGAACCTTGACTTGTATGATCTTCTTGTGCTTCAGACAGTGTCCGTAT 1920  
Db 1861 CCGATTAAACAGAACCTTGACTTGTATGATCTTCTTGTGCTTCAGACAGTGTCCGTAT 1920  
Qy 1921 TTGCGAAAAATTCCGGGTGATGTCAAAGGATCTTCCAAAGATGATCCGCTGAGGCA 1980  
Db 1921 TTGCGAAAAATTCCGGGTGATGTCAAAGGATCTTCCAAAGATGATCCGCTGAGGCA 1980  
Qy 1981 TTGCGAAAAATTCCGGGTGATGTCAAAGGATCTTCCAAAGATGATCCGCTGAGGCA 2040  
Db 1981 TTGCGAAAAATTCCGGGTGATGTCAAAGGATCTTCCAAAGATGATCCGCTGAGGCA 2040  
Qy 2041 TTCTGCTCATGCTTGGCAGCCAGCCAGCAGCTGGGAGCTGAGCCTGGGAGAGCATC 2100  
Db 2041 TTCTGCTCATGCTTGGCAGCCAGCCAGCAGCTGGGAGCTGAGCCTGGGAGAGCATC 2100  
Qy 2101 CTGACAGAGCGGGCCATGAGAACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGGCCATGAGAACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGTGACAGACTCTGAGGAATC 2220  
Db 2161 CAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGTGACAGACTCTGAGGAATC 2220  
Qy 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACCTTGGAGAGGACCACTGAAAGAAAG 2280  
Db 2221 GTCATGGCCAAACCGTAACCTAAGATCCCTCAACCTTGGAGAGGACCACTGAAAGAAAG 2280  
Qy 2281 GATGTAAAGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGTGAGTCTTGAAG 2340  
Db 2281 GATGTAAAGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGTGAGTCTTGAAG 2340  
Qy 2341 CTGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400  
Db 2341 CTGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400  
Qy 2401 TCCCCAGCTGAATCTCTGAGCCTGAGAGGAAACAAAGGTGACAGACCAAGGAGATATG 2460  
Db 2401 TCCCCAGCTGAATCTCTGAGCCTGAGAGGAAACAAAGGTGACAGACCAAGGAGATATG 2460  
Qy 2461 CCTCTCAGTATGCTTGAAGATCTTCCAGTGTGCGCTGACAGACCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTATGCTTGAAGATCTTCCAGTGTGCGCTGACAGACCTGATCTGAGAGAC 2520  
Qy 2521 TGTGGCATCAACGAGCGGTTGCGAGAGTGTGCGCTGACAGACCTGATCTGAGAGAC 2580  
Db 2521 TGTGGCATCAACGAGCGGTTGCGAGAGTGTGCGCTGACAGACCTGATCTGAGAGAC 2580  
Qy 2581 TTGAGACACCTGTGCTATCCAAACAGCCTGGGAGAAACAAAGGTGATATCTATCTGT 2640  
Db 2581 TTGAGACACCTGTGCTATCCAAACAGCCTGGGAGAAACAAAGGTGATATCTATCTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGTAGTCTGAGAGAGGCTGATCTGATCTGAGAGAGAC 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTAGTCTGAGAGAGGCTGATCTGATCTGAGAGAGAC 2700  
Qy 2701 GACAGGCTGTGTGTGTTTCTTTCACCTTGGCTTATGAGTATCTATGCTGAGAGAC 2760  
Db 2701 GACAGGCTGTGTGTGTTTCTTTCACCTTGGCTTATGAGTATCTATGCTGAGAGAC 2760  
Qy 2761 CTGAGCCTTGAAGTAAACCTGTGTGAAGCAATGTGAGTAAAGCTTGTGAGAGAGATG 2820  
Db 2761 CTGAGCCTTGAAGTAAACCTGTGTGAAGCAATGTGAGTAAAGCTTGTGAGAGAGATG 2820  
Qy 2821 AGAGAACCATCTTGTATCTCTGAGAGACCTGAGTGTGTAAGTGTATCTACCCGCGG 2880  
Db 2821 AGAGAACCATCTTGTATCTCTGAGAGACCTGAGTGTGTAAGTGTATCTACCCGCGG 2880

Qy 2881 TGTGTGAGAGTGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940  
Db 2881 TGTGTGAGAGTGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940  
Qy 2941 ACGGACAAATGCTGCTGT 3000  
Db 2941 ACGGACAAATGCTGCTGT 3000  
Qy 3001 AACAGTGTCTGAGAGAGCTGCGGTGAAGGCAATGTGATCTGATCTTGTGTGTGTGT 3060  
Db 3001 AACAGTGTCTGAGAGAGCTGCGGTGAAGGCAATGTGATCTGATCTTGTGTGTGTGT 3060  
Qy 3061 GCACTGTCTGCGCTTGTGCGCAACCGGATGTGACAGTCTTAAACCTGGTGTGAGAT 3120  
Db 3061 GCACTGTCTGCGCTTGTGCGCAACCGGATGTGACAGTCTTAAACCTGGTGTGAGAT 3120  
Qy 3121 AACTTCAGTCCCAAGAGATGATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180  
Db 3121 AACTTCAGTCCCAAGAGATGATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180  
Qy 3181 TTACAGATTAATTTGGCTGTGAAATGTGCAATCTGTGCAATTAAGAGTGTGTGTGT 3240  
Db 3181 TTACAGATTAATTTGGCTGTGAAATGTGCAATCTGTGCAATTAAGAGTGTGTGTGT 3240  
Qy 3241 GAAAGTGTGCTGCTGCTGTGAAATGTGCAATCTGTGCAATTAAGAGTGTGTGTGTGT 3300  
Db 3241 GAAAGTGTGCTGCTGCTGTGAAATGTGCAATCTGTGCAATTAAGAGTGTGTGTGTGT 3300  
Qy 3301 GATGACCGACCAAAATGTGACTTACTTTCGCGCTTCCGTAAGGCGGAGTGTGTGTGT 3360  
Db 3301 GATGACCGACCAAAATGTGACTTACTTTCGCGCTTCCGTAAGGCGGAGTGTGTGTGT 3360  
Qy 3361 GCTTGT 3420  
Db 3361 GCTTGT 3420  
Qy 3421 TTCAAGAGCATTAACCAATTTGCAAGTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGT 3480  
Db 3421 TTCAAGAGCATTAACCAATTTGCAAGTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGT 3480  
Qy 3481 CAGAGAGTGTGACAAAGTGTGACAGAGCTTCCGCAACCCATGTGAGAGCAAGCAAA 3540  
Db 3481 CAGAGAGTGTGACAAAGTGTGACAGAGCTTCCGCAACCCATGTGAGAGCAAGCAAA 3540  
Qy 3541 CAAGATTAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3600  
Db 3541 CAAGATTAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3600  
Qy 3601 GGGCTTGTGATCAACAGTGTGATCAATGACAGAGGATATGAGCTGTGTGTGTGTGT 3660  
Db 3601 GGGCTTGTGATCAACAGTGTGATCAATGACAGAGGATATGAGCTGTGTGTGTGTGT 3660  
Qy 3661 GAGCTGAGCTGAGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3720  
Db 3661 GAGCTGAGCTGAGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3720  
Qy 3721 CACTGTGAGAGGCTTGT 3780  
Db 3721 CACTGTGAGAGGCTTGT 3780  
Qy 3781 GTGTCTGT 3840  
Db 3781 GTGTCTGT 3840  
Qy 3841 CACAGCGGT 3900  
Db 3841 CACAGCGGT 3900  
Qy 3901 GCTGATGACCAAGAGGCTGT 3960  
Db 3901 GCTGATGACCAAGAGGCTGT 3960



Db 1081 CAAGGTTGCGTGCATGAAACAACGTAAGCTGTCGACCAAGTGCAGGTCGCCGC 1140  
Qy 1141 GTGGGCTCTCATCTGCGTGGCCCTGCAAGCTGAGAGTGTGGGGAGAGCTGCGC 1200  
Db 1141 GTGGGCTCTCATCTGCGTGGCCCTGCAAGCTGAGAGTGTGGGGAGAGCTGCGC 1200  
Qy 1201 CCGTTCAACCAACGCTGCAAGGCTGCAAGGCTTTTGTGTTTCAATGAGTCAACCCCT 1260  
Db 1201 CCGTTCAACCAACGCTGCAAGGCTGCAAGGCTTTTGTGTTTCAATGAGTCAACCCCT 1260  
Qy 1261 CGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGATTGCTCTGAAGGCTTCTG 1320  
Db 1261 CGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGATTGCTCTGAAGGCTTCTG 1320  
Qy 1321 CGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGTGAAGCACTCATG 1380  
Db 1321 CGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGTGAAGCACTCATG 1380  
Qy 1381 GTTCAAGGACTGGGGAGCTGAGCTCGTGGCTGTTCACATGAAACATCCCTTCCCA 1440  
Db 1381 GTTCAAGGACTGGGGAGCTGAGCTCGTGGCTGTTCACATGAAACATCCCTTCCCA 1440  
Qy 1441 GACAGCACTGAGAGAGTACTACACTTCTTCACTCAAGTCTCCAGACTTCTGTGC 1500  
Db 1441 GACAGCACTGAGAGAGTACTACACTTCTTCACTCAAGTCTCCAGACTTCTGTGC 1500  
Qy 1501 GCGTTGACTAGTGTGAAGAGGCTGTGAATGAGCCAGCTCTGCTCTGTACGTT 1560  
Db 1501 GCGTTGACTAGTGTGAAGAGGCTGTGAATGAGCCAGCTCTGCTCTGTACGTT 1560  
Qy 1561 GAGAGAGCAAGAGGCTTAAAGAGCTTAAACAGGAGGCTTCATATCACTCGCTTGG 1620  
Db 1561 GAGAGAGCAAGAGGCTTAAAGAGCTTAAACAGGAGGCTTCATATCACTCGCTTGG 1620  
Qy 1621 ATGAAGCGTTTCTGTTGGGCTCTGAGAGCAAGAGTGAAGAGGCACTGAGAGTCTG 1680  
Db 1621 ATGAAGCGTTTCTGTTGGGCTCTGAGAGCAAGAGTGAAGAGGCACTGAGAGTCTG 1680  
Qy 1681 CTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGAGCTTCTGCACTGGGCTCTCTGTTG 1740  
Db 1681 CTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGAGCTTCTGCACTGGGCTCTCTGTTG 1740  
Qy 1741 GGTCAAGAGCTTAATGCCACAACCCAGAGAGCAACCTGAGAGGCTTCACTGTCTTTC 1800  
Db 1741 GGTCAAGAGCTTAATGCCACAACCCAGAGAGCAACCTGAGAGGCTTCACTGTCTTTC 1800  
Qy 1801 GAGACTCAAGACAAAGATTGTTGCTGCTGCAATTAAACAGCTTCCAAAGAGTGTGCTT 1860  
Db 1801 GAGACTCAAGACAAAGATTGTTGCTGCTGCAATTAAACAGCTTCCAAAGAGTGTGCTT 1860  
Qy 1861 CGGATTAAACAGAACTGAGCTTGAATGATCTTCTGCTGCTGCAAGACTGTCCGAT 1920  
Db 1861 CGGATTAAACAGAACTGAGCTTGAATGATCTTCTGCTGCTGCAAGACTGTCCGAT 1920  
Qy 1921 TTGCGGAAAAATTCGGGTGTGATGTCAAAAGGATCTTCCCAAGATGATGCCCTGAGGA 1980  
Db 1921 TTGCGGAAAAATTCGGGTGTGATGTCAAAAGGATCTTCCCAAGATGATGCCCTGAGGA 1980  
Qy 1981 TGTCTGTGGTCCCTCTATGATGAGGAGTAAAGACCTCATTTAGAGAGAGTGTGAGAT 2040  
Db 1981 TGTCTGTGGTCCCTCTATGATGAGGAGTAAAGACCTCATTTAGAGAGAGTGTGAGAT 2040  
Qy 2041 TTCTGCTCATGCTGTGGCAACCAACCACTGCGGCACTGAGACTGTGGGCAAGCAATC 2100  
Db 2041 TTCTGCTCATGCTGTGGCAACCAACCACTGCGGCACTGAGACTGTGGGCAAGCAATC 2100  
Qy 2101 CTGACAGAGCGGCGCATGAAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGCGCATGAAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTTGAATGACAGATTACCCCTGTGTGAGAGCACTCTGAGAGATC 2220  
Db 2161 CAGACCTGATGTTTGAATGACAGATTACCCCTGTGTGAGAGCACTCTGAGAGATC 2220

Qy 2221 GTTATGGCAACCGTAACCTTAAGATCCCTCAACTTTGGAGGCAACCACTGAAGAGAG 2280  
Db 2221 GTTATGGCAACCGTAACCTTAAGATCCCTCAACTTTGGAGGCAACCACTGAAGAGAG 2280  
Qy 2281 GATGTAAAGATGGCTGTGAAGCTTAAACACCCAAATGTTGTGGAGTCTTTGAGG 2340  
Db 2281 GATGTAAAGATGGCTGTGAAGCTTAAACACCCAAATGTTGTGGAGTCTTTGAGG 2340  
Qy 2341 CTGATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
Db 2341 CTGATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
Qy 2401 TCCCCAGCTGAAATCTCTGAGCTGTGAGAGGAAACAGGTGACAGACAGGAGTAAATG 2460  
Db 2401 TCCCCAGCTGAAATCTCTGAGCTGTGAGAGGAAACAGGTGACAGACAGGAGTAAATG 2460  
Qy 2461 CCTCTCAAGTATGCTTGAAGAGTCTCCAGATGCGCTGAGAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAAGTATGCTTGAAGAGTCTCCAGATGCGCTGAGAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGGCATCAAGCCAGGAGTTCAGAGTCTGAGAGCTTCTGAGAGCAACCGAGAC 2580  
Db 2521 TGTGGCATCAAGCCAGGAGTTCAGAGTCTGAGAGCTTCTGAGAGCAACCGAGAC 2580  
Qy 2581 TTGACACACTGTGTGCTTATCCAAACAAGCTGTGGGAAACGAAGTGTAAATCTACTGT 2640  
Db 2581 TTGACACACTGTGTGCTTATCCAAACAAGCTGTGGGAAACGAAGTGTAAATCTACTGT 2640  
Qy 2641 CGATCCATGAGGCTTCCCACTGTAGTGTGAGAGGCTGATGTCAATCAAGTCCACTG 2700  
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTGTGAGAGGCTGATGTCAATCAAGTCCACTG 2700  
Qy 2701 GACAGGCTGTGTGTTTTCTTTCACCTTGCGCTTAATGGTAACTCATAGCTGACAC 2760  
Db 2701 GACAGGCTGTGTGTTTTCTTTCACCTTGCGCTTAATGGTAACTCATAGCTGACAC 2760  
Qy 2761 CTGAGCTTATGATTAACCTGTGTGAGAGCAATGAGGCTTCTGTGTGAGGTCATG 2820  
Db 2761 CTGAGCTTATGATTAACCTGTGTGAGAGCAATGAGGCTTCTGTGTGAGGTCATG 2820  
Qy 2821 AGAGAACCATCTTGCATCTCCAGAGACTGTGAGTGTAAAGTGCATCTCACGCCGCG 2880  
Db 2821 AGAGAACCATCTTGCATCTCCAGAGACTGTGAGTGTAAAGTGCATCTCACGCCGCG 2880  
Qy 2881 TGTGTGAGAGTGTGTCTGTGTGATCTGAGAGACAGACCTGAAAGGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTGTGTCTGTGTGATCTGAGAGACAGACCTGAAAGGCTGTGATCTC 2940  
Qy 2941 ACGACAAATGCCCTGGGTGAGAGGTGGGTGTGCTGCTGTGAGAGAGCTGAGAGAAAG 3000  
Db 2941 ACGACAAATGCCCTGGGTGAGAGGTGGGTGTGCTGCTGTGAGAGAGCTGAGAGAAAG 3000  
Qy 3001 AACAGTGTGAGAGAGCTGGGGTTGAAGGATGTGACCTGATGATGATGATGATGATGAT 3060  
Db 3001 AACAGTGTGAGAGAGCTGGGGTTGAAGGATGTGACCTGATGATGATGATGATGATGAT 3060  
Qy 3061 GCACTGTCTTGGCCCTTCTGCAACCGGATCTGACAGCTTAAACCTGTGTGAGAT 3120  
Db 3061 GCACTGTCTTGGCCCTTCTGCAACCGGATCTGACAGCTTAAACCTGTGTGAGAT 3120  
Qy 3121 AACTTCACTCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180  
Db 3121 AACTTCACTCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180  
Qy 3181 TTACAGATTAATGGCTGTGAGAAATGAGATGATGATGATGATGATGATGATGATGATGAT 3240  
Db 3181 TTACAGATTAATGGCTGTGAGAAATGAGATGATGATGATGATGATGATGATGATGATGAT 3240  
Qy 3241 GAAAGTGAAGTACTCAAGCCCGAGTCTGAATGATGATGATGATGATGATGATGATGATGAT 3300  
Db 3241 GAAAGTGAAGTACTCAAGCCCGAGTCTGAATGATGATGATGATGATGATGATGATGATGAT 3300

```

QY 3301 GATGACCGACCAAAATAGACTTATCTTCCGGCTCCCTGAAAGCCGGGCAATGACCAGT 3360
DB 3301 GATGACCGACCAAAATAGACTTATCTTCCGGCTCCCTGAAAGCCGGGCAATGACCAGT 3360
QY 3361 GCCTTGCTGTGGGGATGAACTCCAGACGAAGAACCTGTGTCCCTTCTGCTGGAAGAC 3420
DB 3361 GCCTTGCTGTGGGGATGAACTCCAGACGAAGAACCTGTGTCCCTTCTGCTGGAAGAC 3420
QY 3421 TTCAAGAGAGATGACGATTTTCCAAAGTCTCTCTGCTGGGCCCAATGATGATCC 3480
DB 3421 TTCAAGAGAGATGACGATTTTCCAAAGTCTCTCTGCTGGGCCCAATGATGATCC 3480
QY 3481 CAGAGAGTTGACCAAGTGGAGAGAGCTCCCGCAACCCATGAGGAGCAAGCAACAA 3540
DB 3481 CAGAGAGTTGACCAAGTGGAGAGAGCTCCCGCAACCCATGAGGAGCAAGCAACAA 3540
QY 3541 CAAAGTAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
DB 3541 CAAAGTAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
QY 3601 GGGCTTGGATCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
DB 3601 GGGCTTGGATCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
QY 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGATGATGATGATGATGATGATGAT 3720
DB 3661 GAGCTGAGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGATGATGATGATGATGATGATGAT 3720
QY 3721 CACTGGGAGCGGCTGGGCTTGAAGGGCTGTGTCTTTAAACAGTGTGATGATGATGATGATGATGAT 3780
DB 3721 CACTGGGAGCGGCTGGGCTTGAAGGGCTGTGTCTTTAAACAGTGTGATGATGATGATGATGAT 3780
QY 3781 GGTGCTGAGTCACTGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGATGAT 3840
DB 3781 GGTGCTGAGTCACTGGAGCGGCTGGGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGATGAT 3840
QY 3841 CACAGCGGTGTGCTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900
DB 3841 CACAGCGGTGTGCTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900
QY 3901 GGTGATGACACAGCGGTGTGTCTGTGTCTGATGATGATGATGATGATGATGATGATGATGAT 3960
DB 3901 GGTGATGACACAGCGGTGTGTCTGTGTCTGATGATGATGATGATGATGATGATGATGATGAT 3960
QY 3961 GCTGATGACACAGCGGTGTGTCTGTGTCTGATGATGATGATGATGATGATGATGATGATGAT 4020
DB 3961 GCTGATGACACAGCGGTGTGTCTGTGTCTGATGATGATGATGATGATGATGATGATGATGAT 4020
QY 4021 GGGCTGTGTCTTTAA 4035
DB 4021 GGGCTGTGTCTTTAA 4035

```

## RESULT 3

ABK48628 standard; cDNA; 3900 BP.

ABK48628;

13-AUG-2002 (first entry)

Human MATER cDNA.

Human, gene; ss; contraceptive; antiinfertility; MATER; maternal antigen that embryos require; MATER null phenotype; oocyte; early embryonic survival; premature ovarian failure; POF; autoimmune infertility; chromosome 19; gene therapy; fertility.

Homo sapiens.

Key Location/Qualifiers  
FT 1..3603  
CDS /\*tag= a

```

FT /product= "Human MATER"
XX
XX MO200232955-A1.
XX
XX 25-APR-2002.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX 18-OCT-2000; 2000US-0241510P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX MPI; 2002-454595/48.
XX
XX P-PSDB; AAU79526.
XX
XX New isolated human Maternal Antigen That Embryos Require protein and
XX polynucleotide, useful in diagnosing or treating fertility or reduced
XX fertility, or as a contraceptive.
XX
XX Claim 11; Page 80-85; 93pp; English.
XX
XX The invention discloses an isolated human MATER (Maternal Antigen That
XX Embryos Require) protein, which can complement a MATER null phenotype in
XX which zygotes arising from the oocyte do not progress beyond the two-cell
XX stage. MATER is required for early embryonic survival and abnormal levels
XX of the protein can lead to premature ovarian failure (POF) and can be
XX caused by under or over expression of MATER or an autoimmune response to
XX MATER. MATER is a single-copy maternal effect gene found on chromosome
XX 19. The MATER protein and polynucleotide, by gene therapy, are useful in
XX diagnosing or treating fertility and reduced fertility. In particular,
XX the MATER protein is useful as a contraceptive agent, or for influencing
XX (either inhibiting or enhancing) fertility and can be used to detect a
XX predisposition to infertility or reduced fertility, or for presymptomatic
XX screening of an individual for infertility/reduced fertility. The protein
XX and polynucleotide are also useful for detecting an excess or deficiency,
XX or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
XX human or a mouse) or for screening for a compound useful in influencing
XX MATER-mediated fertility. The sequence presented is the human MATER cDNA.
XX Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to
XX published human genomic DNA sequences, while the remainder of the
XX sequence was determined by direct cloning of human ovarian cDNAs
XX
XX
XX Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;
XX
XX Query Match 78.5%; Score 3166.2; DB 6; Length 3900;
XX Best Local Similarity 96.2%; Pred. No. 0;
XX Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;
XX
QY 1 ATGGAAGAGACAATGCGCTCACCTTTTCCAGCTAGCGGGCTGCAATGATGATGATGATGATGATGAT 60
DB 154 ATGGAAGAGACAATGCGCTCACCTTTTCCAGCTAGCGGGCTGCAATGATGATGATGATGATGATGAT 213
QY 61 CTAGCAAGAGAAATTTTCAAGACATTCAGAGAAATTAAGAAATCTTCAGAAATCG 120
DB 214 CTAGCAAGAGAAATTTTCAAGACATTCAGAGAAATTAAGAAATCTTCAGAAATCG 273
QY 121 ACCACATGCTCTATTTCCACAGCTTTGAAATCGAATGCAACGCTGGAATGTCTGCACTC 180
DB 274 ACCACATGCTCTATTTCCACAGCTTTGAAATCGAATGCAACGCTGGAATGTCTGCACTC 333
QY 181 CTCTTGATGATGATTTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 334 CTCTTGATGATGATTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 393
QY 241 AACATGAACCTGCAGAACCTCTCGAGAAAGCAGGAGATGATGATGATGATGATGATGATGATGAT 289
DB 394 AACATGAACCTGCAGAACCTCTCGAGAAAGCAGGAGATGATGATGATGATGATGATGATGATGAT 453
QY 290 ----- 289
DB 454 GAAGATCCTGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 513

```

[illegible]

QY	1318	GGCCGTAAGGCTGCGGAGGGAGTGTGGAAATGGAAGTCAGCTTTGATGGTGAACACTC	13777
Db	1554	TGCGGTAATGCGTGTGAGAGGGAGTGTGGAAATGGAAGTCAAGTGTTTGATGGTGAACACTC	16533
QY	1378	ATGGTTCAAGACTCGGGAGTCTGAGTCTCCGTCTCTGTTCATCATGAACATCCTTCTC	14379
Db	1654	ATGGTTCAAGGACTCGGGAGTCTGAGTCTCCGTCTCTGTTCATCATGAACATCCTTCTC	17133
QY	1438	CCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACTCACTGCTTCAGAGACTTCTGT	1497
Db	1714	CCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACTCACTGCTTCAGAGACTTCTGT	17733
QY	1498	GCCGCTTGTACTGCTGTGTAAGAGGGCTGGAAATCGAGCCAGGCTCTGCGCTGTGAC	1557
Db	1774	GCCGCTTGTACTGCTGTGTAAGAGGGCTGGAAATCGAGCCAGGCTCTCTGCGCTGTGAC	18333
QY	1558	GTTGAGAGACAAAGAAGTCCATGAGCTTTTAAACAGCAGGCTTCCATATCACTCGCTT	1617
Db	1834	GTTGAGAGAGACAAAGAAGTCCATGAGCTTTTAAACAGCAGGCTTCCATATCACTCGCTT	18933
QY	1618	TGATGAAGCGTTTCTGTGTTTGGCTCTGTGAGCAAGACGTAAAGAGGCCCTGAGAGTCTC	16777
Db	1894	TGATGAAGCGTTTCTGTGTTTGGCTCTGTGAGCAAGACGTAAAGAGGCCCTGAGAGTCTC	19533
QY	1678	CTGCTGGGCTGTCCCGTTCCTCCGTGGGGTGAAGCAGAAAGCTTCTGCACTGGGTCTCTCTG	17377
Db	1954	CTGCTGGGCTGTCCCGTTCCTCCGTGGGGTGAAGCAGAAAGCTTCTGCACTGGGTCTCTCTG	20133
QY	1738	TTGGGTCAGAGCCCTAATGCCACACCCAGAGACACCCCTGAGACGCTTCACTGCTCTT	17977
Db	2014	TTGGGTCAGAGCCCTAATGCCACACCCAGAGACACCCCTGAGACGCTTCACTGCTCTT	20733
QY	1798	TTGAGACTCAGACAAAGAAGTGTTCGCTTGGCATTTAAACAGCTTCCAAAGAGTGTG	18577
Db	2074	TTGAGACTCAGACAAAGAAGTGTTCGCTTGGCATTTAAACAGCTTCCAAAGAGTGTG	21333
QY	1858	CTTCCGATTAACCGAAACTGGAATTGATGACATCTTCTCTGCTTCAGACTGTGCTCG	19177
Db	2134	CTTCCGATTAACCGAAACTGGAATTGATGACATCTTCTCTGCTTCAGACTGTGCTCG	21933
QY	1918	TATTTGGGGAATAATCGGGTGGAGTGCAAAAGGATCTTCCCAAGATAGGTCCGCTGAG	19777
Db	2194	TATTTGGGGAATAATCGGGTGGAGTGCAAAAGGATCTTCCCAAGATAGGTCCGCTGAG	22533
QY	1978	GCATGTCCTGTGTCTCCCTATATGAGATCGGGATTAAGACCTTCATTGAGAGCACTGGGAA	20377
Db	2254	GCATGTCCTGTGTCTCCCTATATGAGATCGGGATTAAGACCTTCATTGAGAGCACTGGGAA	22133
QY	2038	GATTTCTGCTCCATGCTTGGACACCCACCCACACTGTCGGCAGCTGAGACTGGGACAGC	20977
Db	2314	GATTTCTGCTCCATGCTTGGACACCCACCCACACTGTCGGCAGCTGAGACTGGGACAGC	23733
QY	2098	ATCTGTGACAGGGGGGCAATGAACACCTGTGTGCCAAGCTGAGGCAATCCCACTGGACAG	21577
Db	2374	ATCTGTGACAGGGGGGCAATGAACACCTGTGTGCCAAGCTGAGGCAATCCCACTGGACAG	24333
QY	2158	ATACAGACCCCTGATGTTTAAAGATGCAAGATTAACCCCTGGTGTGACAGACTCTGAGAGA	22177
Db	2434	ATACAGACCCCTGATGTTTAAAGATGCAAGATTAACCCCTGGTGTGACAGACTCTGAGAGA	24933
QY	2218	ATGCTCATGGCCACCGTAACTTAAGATCCCTCACTTGGAGAGGCAACCACTTGAGAGAA	22777
Db	2494	ATGCTCATGGCCACCGTAACTTAAGATCCCTCACTTGGAGAGGCAACCACTTGAGAGAA	25533
QY	2278	GAGGATGTAAGGATGGCGTGTGAAGCCTTTAAACACCCAAATATTTGTTGGAGTCTTTTG	23377
Db	2554	GAGGATGTAAGGATGGCGTGTGAAGCCTTTAAACACCCAAATATTTGTTGGAGTCTTTTG	26133
QY	2338	AGGCTGAGTTGCTGTGAGTTGAACCCATGACCTGTTAACCTGAAGATCTCCCAATCTTAAAG	23977
Db	2614	AGGCTGAGTTGCTGTGAGTTGAACCCATGACCTGTTAACCTGAAGATCTCCCAATCTTAAAG	26733
QY	2398	ACCTTCCCCAGCCTGAATCTCTGAGCCTGGCAGAGAAACAAGGTGACAGACAGGAGTA	24577



```

Db 2674 ACCCTCCCCAGCTGAATCTCTGAGCCTGGCAGGAAACAAGTACAGACAGGAGATA 2733
QY 2458 ATGCTCTCAGTGAAGCCTTGAAGAGTCTCCAGTGGCGCCCTGCAAACTGATCTGAG 2517
Db 2734 ACGCTCTCAGTGAAGCCTTGAAGAGTCTCCAGTGGCGCCCTGCAAACTGATCTGAG 2793
QY 2518 GACTGTGCAATCAAGCAAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACCG 2577
Db 2794 GACTGTGCAATCAAGCAAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACCG 2853
QY 2578 AGCTTGAACAACCTGTGCTTATCCAAACAAGCTGGGAAAGAAAGTGTAAATCTACTG 2637
Db 2854 AGCTTGAACAACCTGTGCTTATCCAAACAAGCTGGGAAAGAAAGTGTAAATCTACTG 2913
QY 2638 TGTGATTCATGAGGCTTCCCACTGTAGTCTGCAAGAGGCTGATGCTGAATCAGTCCAC 2697
Db 2914 TGTGATTCATGAGGCTTCCCACTGTAGTCTGCAAGAGGCTGATGCTGAATCAGTCCAC 2973
QY 2698 CTGCAACAGGCTGCTGTGTGTTTCTTCACTTGGCGTTAAGGTAACATGAGCTGACG 2757
Db 2974 CTGCAACAGGCTGCTGTGTGTTTCTTCACTTGGCGTTAAGGTAACATGAGCTGACG 3033
QY 2758 CACTGAGCCTTGAATGAACCTCTGTGAAGAACAATGCGTGAACCTTCTGTGCAAGCTC 2817
Db 3034 CACTGAGCCTTGAATGAACCTCTGTGAAGAACAATGCGTGAACCTTCTGTGCAAGCTC 3093
QY 2818 ATGAGAGAACAATCTGTGATCTCCAGAACCTGGAGTTGTAAAGTGTATCTCACTCAGCC 2877
Db 3094 ATGAGAGAACAATCTGTGATCTCCAGAACCTGGAGTTGTAAAGTGTATCTCACTCAGCC 3153
QY 2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTCGAAGACAGACACCTGAAGAGCTGAT 2937
Db 3154 GCGTCTGTGAGAGTCTGTCTGTGTGATCTCGAAGACAGACACCTGAAGAGCTGAT 3213
QY 2938 CTCACGGAACAATGCTTGGGTGAACGTTGGGTTGCTGCGCTGTGCAAGAGCTGAAGCAA 2997
Db 3214 CTCACGGAACAATGCTTGGGTGAACGTTGGGTTGCTGCGCTGTGCAAGAGCTGAAGCAA 3273
QY 2998 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGATGAGCATGACTTCTGATTTGCTGT 3057
Db 3274 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGATGAGCATGACTTCTGATTTGCTGT 3333
QY 3058 GAGGACTCTCTCTTGGCCCTTCTGCAACCGGCACTGACCAAGTCTAAACCTGGTGAG 3117
Db 3334 GAGGACTCTCTCTTGGCCCTTCTGCAACCGGCACTGACCAAGTCTAAACCTGGTGAG 3393
QY 3118 AATTAATTCACTCCCAAGAAATGATGAAGCTGTGTGGCTTGTGCTGTCCAGCTCT 3177
Db 3394 AATTAATTCACTCCCAAGAAATGATGAAGCTGTGTGGCTTGTGCTGTCCAGCTCT 3453
QY 3178 AACTTAAGATTAATTGGGCTGTGGAATGGCAGTACCTGTGCAAAATTAAGAAAGTGTG 3237
Db 3454 AACTTAAGATTAATTGGGCTGTGGAATGGCAGTACCTGTGCAAAATTAAGAAAGTGTG 3513
QY 3238 GAGGAAGTGAAGTCTGAAGCCCGAGTCTGAATTTGACGAGTGTGGCATTTCTTTGAT 3297
Db 3514 GAGGAAGTGAAGTCTGAAGCCCGAGTCTGAATTTGACGAGTGTGGCATTTCTTTGAT 3573
QY 3298 GAAGATGAACGACAC 3312
Db 3574 GAAGATGAACGAGTAC 3588

```

RESULT 4  
AAD49018  
ID AAD49018 standard; cDNA; 3900 BP.

XX AAD49018;  
XX 07-MAR-2003 (first entry)  
XX Human MATER cDNA.

```

XX XX Human; MATER protein; infertility; fertility; contraceptive agent;
KW KW gene therapy; gene; ss.
XX OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..3603
FT /tag= a
FT /product= "Human MATER protein"
XX
XX WO200281492-A1.
XX
XX 17-OCT-2002.
XX
XX 29-MAR-2002; 2002WO-US009776.
XX
XX 04-APR-2001; 2001WO-US010981.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson LM, Tong Z;
XX
XX WPI; 2003-058494/05.
XX
XX P-PSDB; AAE31749.
XX
XX
XX New isolated variant MATER proteins and nucleic acids, useful for
XX PT diagnosing, prognosing and treating infertility and reduced fertility,
XX PT and as contraceptive agents.
XX
XX Example 2; Page 93-98; 110pp; English.
XX
XX The present invention relates to novel MATER proteins and polynucleotides
XX encoding such proteins. The MATER proteins are essential to fertility.
XX Sequences of the invention are useful for diagnosing, prognosing and
XX treating infertility, reduced fertility and as contraceptive agents. They
XX are also useful in gene therapy. The method is useful for detecting a
XX predisposition to or pre-symptomatic screening of an individual for
XX infertility or reduced fertility. The present sequence is human MATER
XX CC cDNA
XX
XX Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;
SQ
XX
XX Query Match 78.5%; Score 3166.2; DB 8; Length 3900;
XX Best Local Similarity 96.2%; Pred. No. 0;
XX Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;
QY 1 ATGGAAGGAGCAAAATGCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 60
Db 154 ATGGAAGGAGCAAAATGCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 213
QY 61 CTAAACAAGGAAGAAATTCAGACATTCAGAGAAATTAAGAAAGAAATCTTCAGAAATCG 120
Db 214 CTAAACAAGGAAGAAATTCAGACATTCAGAGAAATTAAGAAAGAAATCTTCAGAAATCG 273
QY 121 ACCACATGCTTATTCACAGTTTGAATCGAAGATCCAAACGTGAATGTCTGGCACTC 180
Db 274 ACCACATGCTTATTCACAGTTTGAATCGAAGATCCAAACGTGAATGTCTGGCACTC 333
QY 181 CTCTTGATGAGATTAATGAGATCGCTGGCTGGCTGACATTCATTAAGATCTTTGAA 240
Db 334 CTCTTGATGAGATTAATGAGATCGCTGGCTGGCTGACATTCATTAAGATCTTTGAA 393
QY 241 AACATGAACCTGGAGAACCTCTCGGAAGAGGCAAGGATGACATGAAGAA----- 289
Db 394 AACATGAACCTGGAGAACCTCTCGGAAGAGGCAAGGATGACATGAAGAAAGACATTACCA 453
QY 290 ----- 289
Db 454 GAAGATCTGAAGCAAGATGACTGACCAAGAACCAAGAAAGAAAGTCCAGAAAT 513
QY 290 -----AAATTTCA 297

```



Db 514 TCACAGCTGTGCACACAGATAGTCCACAGCTGCAGAGACAAAGAAACAGAAATTTCAC 573  
Qy 298 CAAGCTATGGAACAAAGAGGTGCACAGAGACAGAGACAAAGAAATTTCACAA 357  
Db 574 CAAGCTATGGAACAAAGAGGTGCACAGAGACAGAGACAAAGAAATTTCACAA 633  
Qy 358 GCTATGGAACAAAGAGGTGCACAGAGACAGAGACAAAGAAATTTCACAA 417  
Db 634 GCTATGGAACAAAGAGGTGCACAGAGACAGAGACAAAGAAATTTCACAA 693  
Qy 418 ACATGGAGCTACAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGAGATGTACCTGCT 477  
Db 694 ACATGGAGCTACAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGAGATGTACCTGCT 753  
Qy 478 AATTTTGAACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537  
Db 754 AATTTTGAACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813  
Qy 538 TCAGACCCGGTGGGGCTTCCGGGCTGCACAGTGTCTGTCACGAAAGTCAGAAATTGGG 597  
Db 814 TCAGACCCGGTGGGGCTTCCGGGCTGCACAGTGTCTGTCACGAAAGTCAGAAATTGGG 873  
Qy 598 AAATGAGCTATGACCAAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657  
Db 874 AAATGAGCTATGACCAAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 933  
Qy 658 AATGTTCTCTACGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 717  
Db 934 AATGTTCTCTACGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 993  
Qy 718 GTACAGAGTTTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGGTGCAGAGATCATG 777  
Db 994 GTACAGAGTTTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGGTGCAGAGATCATG 1053  
Qy 778 TCCGACCAAGAAAGCTGT 837  
Db 1054 TCCGACCAAGAAAGCTGT 1113  
Qy 838 AACATGACACAAAGCTGTGCACAAAGCTGTGGCTGAGAGACCTTCCTTCACCTCATG 897  
Db 1114 AACATGACACAAAGCTGTGCACAAAGCTGTGGCTGAGAGACCTTCCTTCACCTCATG 1173  
Qy 898 CGCAGTCTGCTGAGGAAGTCTGCTCCCTGAGTCCCTTCATGATGTCACGTCAGAGAC 957  
Db 1174 CGCAGTCTGCTGAGGAAGTCTGCTCCCTGAGTCCCTTCATGATGTCACGTCAGAGAC 1233  
Qy 958 GTGGCACAAGAGAGCTCAAGTCAAGAGTCTGTCTCCCTTACCTGTATGATGAGGA 1017  
Db 1234 GTGGCACAAGAGAGCTCAAGTCAAGAGTCTGTCTCCCTTACCTGTATGATGAGGA 1293  
Qy 1018 ATCTCCGGGGAACAAAGATCCATTGCTCTTGAAGCCGGGATTTGTGAGCATCAAG 1077  
Db 1294 ATCTCCGGGGAACAAAGATCCATTGCTCTTGAAGCCGGGATTTGTGAGCATCAAG 1353  
Qy 1078 ACACAAAGGTTGCGGCGATCAAGAACACCGTGAGCTGTCGACAGTCCAGGTTGCC 1137  
Db 1354 ACACAAAGGTTGCGGCGATCAAGAACACCGTGAGCTGTCGACAGTCCAGGTTGCC 1413  
Qy 1138 GCGGTGAGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197  
Db 1414 GCGGTGAGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473  
Qy 1198 GCGGCTTTCACAAACGCTCAAGGCTGACGCGCTTTTGTGTTCATCAAGTCAAC 1257  
Db 1474 GCGGCTTTCACAAACGCTCAAGGCTGACGCGCTTTTGTGTTCATCAAGTCAAC 1533  
Qy 1258 CCTGAGGCGTGTCCGCGCTGCTCAATCTGAGGAAAGGCTGCTGAGGCTTC 1317  
Db 1534 CCTGAGGCGTGTCCGCGCTGCTCAATCTGAGGAAAGGCTGCTGAGGCTTC 1593  
Qy 1318 TGCCGTATGTGCTGTGAGAGAGTGTGAATAGGAAGTCAAGTGTGTGATGATGATGATGAT 1377  
Db 1594 TGCCGTATGTGCTGTGAGAGAGTGTGAATAGGAAGTCAAGTGTGTGATGATGATGATGAT 1653

Qy 1378 AATGTTCAAGAGCTGCGGGAGTCTGAGCTCCGCTGCTGTTTCAATGAAATCTTCTC 1437  
Db 1654 AATGTTCAAGAGCTGCGGGAGTCTGAGCTCCGCTGCTGTTTCAATGAAATCTTCTC 1713  
Qy 1438 CCAGACAGCACTGAGAGAGTACTACCTCTTCTTCCACTGCTCAAGTCTCAAGTCTTCTGT 1497  
Db 1714 CCAGACAGCACTGAGAGAGTACTACCTCTTCTTCCACTGCTCAAGTCTCAAGTCTTCTGT 1773  
Qy 1498 GCGGCTTGTACTACGTGTGAGAGGCTTGAATGAGCCAGCTCTGCTCTGTAC 1557  
Db 1774 GCGGCTTGTACTACGTGTGAGAGGCTTGAATGAGCCAGCTCTGCTCTGTAC 1833  
Qy 1558 GTTGAAGAACAAAGAGTCCATGAGAGCTTAAACAGAGAGCTTCAATGCACTGCTT 1617  
Db 1834 GTTGAAGAACAAAGAGTCCATGAGAGCTTAAACAGAGAGCTTCAATGCACTGCTT 1893  
Qy 1618 TGATGAAGAGCTTCTTGTGTGCTCTGAGAGCAAGTAAAGAGGCTCACTGAGGTC 1677  
Db 1894 TGATGAAGAGCTTCTTGTGTGCTCTGAGAGCAAGTAAAGAGGCTCACTGAGGTC 1953  
Qy 1678 CTGCTGGGCTGTCCCGTTCCTGCGGGGTGAAGCAAGGCTTTCAGCTGGGTCTCTG 1737  
Db 1954 CTGCTGGGCTGTCCCGTTCCTGCGGGGTGAAGCAAGGCTTTCAGCTGGGTCTCTG 2013  
Qy 1738 TTGGGTGACAGCCCTTAATGCCACCACTCCAGAGACACCTCTGAGAGCTTCTCACTGTCTT 1797  
Db 2014 TTGGGTGACAGCCCTTAATGCCACCACTCCAGAGACACCTCTGAGAGCTTCTCACTGTCTT 2073  
Qy 1798 TTGCAAGCTCAAGAACAAAGAGTGTGTGCTTGTGCAATTAACAGCTTCAAGAAAGTGTG 1857  
Db 2074 TTGCAAGCTCAAGAACAAAGAGTGTGTGCTTGTGCAATTAACAGCTTCAAGAAAGTGTG 2133  
Qy 1858 CTTCGATTAACAGAACCTGACCTGTATGATCTTCTTCTGCTCAGCACTGTCTG 1917  
Db 2134 CTTCGATTAACAGAACCTGACCTGTATGATCTTCTTCTGCTCAGCACTGTCTG 2193  
Qy 1918 TATTTGCGGAAATTCGGGTGAGATGTCAAAAGGATCTTCCCAAGATGATGATGATGATGAT 1977  
Db 2194 TATTTGCGGAAATTCGGGTGAGATGTCAAAAGGATCTTCCCAAGATGATGATGATGATGAT 2253  
Qy 1978 GCATGCTCTGTGTGCTCTATGAGTGTGAGGATTAAGACCTCATTTAGAGAGAGTGGGA 2037  
Db 2254 GCATGCTCTGTGTGCTCTATGAGTGTGAGGATTAAGACCTCATTTAGAGAGAGTGGGA 2113  
Qy 2038 GATTTCTGCTCATGCTTGTGACCACTCAAGCTGCGGCACTGAGACCTGCGGACGAC 2097  
Db 2314 GATTTCTGCTCATGCTTGTGACCACTCAAGCTGCGGCACTGAGACCTGCGGACGAC 2373  
Qy 2098 ATCTGACAGAGCGGGCATGAAGACCTGTGTGTCMAAGCTGAGGATCCACTGCAAG 2157  
Db 2374 ATCTGACAGAGCGGGCATGAAGACCTGTGTGTCMAAGCTGAGGATCCACTGCAAG 2433  
Qy 2158 ATACAGACCTGATGTTAAGAAATGCAAGATTACCTCTGCTGTGAGCACTCTGAGGA 2217  
Db 2434 ATACAGACCTGATGTTAAGAAATGCAAGATTACCTCTGCTGTGAGCACTCTGAGGA 2493  
Qy 2218 ATCTCATGAGCAACCTGTAACCTTAAGATCCCTCAACTTGTGAGGACCACTGAAGAA 2277  
Db 2494 ATCTCATGAGCAACCTGTAACCTTAAGATCCCTCAACTTGTGAGGACCACTGAAGAA 2553  
Qy 2278 GAGGATGTAAGAGAGGCTGTGAAGCTTAAACCAACCAAAATGTTGTGAGTCTTGTG 2337  
Db 2554 GAGGATGTAAGAGAGGCTGTGAAGCTTAAACCAACCAAAATGTTGTGAGTCTTGTG 2613  
Qy 2338 AGGCTGATGCTGTGATGTAACCATGCTGTATCACTGAAGATCTCCAAATCTTAAG 2397  
Db 2614 AGGCTGATGCTGTGATGTAACCATGCTGTATCACTGAAGATCTCCAAATCTTAAG 2673  
Qy 2398 ACTTCCCAAGCTGAAGATCTGTAGGCTGACAGAAACAAAGTGAAGACCAAGGAGTA 2457  
Db 2674 ACTTCCCAAGCTGAAGATCTGTAGGCTGACAGAAACAAAGTGAAGACCAAGGAGTA 2733

QY 2458 ATGCTCTCAGTATGCTTGGAGAGTCTCCAGTGCCTCGACAGAGCTGATCTGAG 2517  
 DB 2734 ACGCTCTCAGTATGCTTGGAGAGTCTCCAGTGCCTCGACAGAGCTGATCTGAG 2793  
 QY 2518 GACTGTGATCAACAGCCAGCGGTTGCCAGAGTCTGGCTTCAGCCCTGTCAGCAACGG 2577  
 DB 2794 GACTGTGATCAACAGCCAGCGGTTGCCAGAGTCTGGCTTCAGCCCTGTCAGCAACGG 2853  
 QY 2578 AGCTTGACACACCTTGGCTTATCCAAACAAGCCGGGGAAAGAGAGTAAATCTACTG 2637  
 DB 2854 AGCTTGACACACCTTGGCTTATCCAAACAAGCCGGGGAAAGAGAGTAAATCTACTG 2913  
 QY 2638 TGTGATCATAGAGGCTTCCCACTGTAGTCTGACAGAGCTGATCTGATCATAGTCCAC 2697  
 DB 2914 TGTGATCATAGAGGCTTCCCACTGTAGTCTGACAGAGCTGATCTGATCATAGTCCAC 2973  
 QY 2698 CTGACACGGCTGGCTGTGGTTTCTTGCACTTGGCTTAAAGGTAATCATGGCTGACG 2757  
 DB 2974 CTGACACGGCTGGCTGTGGTTTCTTGCACTTGGCTTAAAGGTAATCATGGCTGACG 3033  
 QY 2758 CACCTGAGCTTAGATGAACCTGAGAAAGCAATGAGCGTGAAGCTTCTGCGAGGTC 2817  
 DB 3034 CACCTGAGCTTAGATGAACCTGAGAAAGCAATGAGCGTGAAGCTTCTGCGAGGTC 3093  
 QY 2818 ATGAGAGAACCATCTTGTCAATCTCCAGGACCTGAGTTGTAAAGTCAATCTCACCGCC 2877  
 DB 3094 ATGAGAGAACCATCTTGTCAATCTCCAGGACCTGAGTTGTAAAGTCAATCTCACCGCC 3153  
 QY 2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGACAGACACCTGAAGAGCTTGAT 2937  
 DB 3154 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGACAGACACCTGAAGAGCTTGAT 3213  
 QY 2938 CTCACGAGCAATATGCTGGGTGAAGGTTGGTGGCTGTGAGAGGACTGAAGCA 2997  
 DB 3214 CTCACGAGCAATATGCTGGGTGAAGGTTGGTGGCTGTGAGAGGACTGAAGCA 3273  
 QY 2998 AAGAACAGTGTCTGACGAGACTCGGTTGAAGGAGTGTGACTGACTTGTGATTTGCTGT 3057  
 DB 3274 AAGAACAGTGTCTGACGAGACTCGGTTGAAGGAGTGTGACTGACTTGTGATTTGCTGT 3333  
 QY 3058 GAGGACATCTCTTGGCCCTTCTGCAACCGGCAATCTGACAGTCTAAACCTGGTGCAG 3117  
 DB 3334 GAGGACATCTCTTGGCCCTTCTGCAACCGGCAATCTGACAGTCTAAACCTGGTGCAG 3393  
 QY 3118 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGGCCCTTGGCTGCCACGCT 3177  
 DB 3394 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGGCCCTTGGCTGCCACGCT 3453  
 QY 3178 AACTTACAGATTAATGGCTGTGAATGGAGTACCTGTGCAATTAAGAGAGCTGCTG 3237  
 DB 3454 AACTTACAGATTAATGGCTGTGAATGGAGTACCTGTGCAATTAAGAGAGCTGCTG 3513  
 QY 3238 GAGGAAGTGCAGTACTGAAGCCCGGAGTCTGAATTTGACGGTAGTTGCAATCTTTTAT 3297  
 DB 3514 GAGGAAGTGCAGTACTGAAGCCCGGAGTCTGAATTTGACGGTAGTTGCAATCTTTTAT 3573  
 QY 3298 GAAGATGACCGACAC 3312  
 DB 3574 GAAGATGACCGGTAC 3588

RESULT 5  
 ADA45220  
 ID ADA45220 standard; DNA; 3830 BP.

AC ADA45220;  
 XX 20-NOV-2003 (first entry)

DE Human MATER splice variant DNA SEQ ID 3.

KW human; MATER; maternal antigen that embryos require; NTP-ase;  
 nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;

KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;  
 KW ovarian dysfunction; autoimmune premature ovarian failure;  
 KW autoimmune disease; contraceptive; autoantibody; ds; gene.  
 OS Homo sapiens.  
 PH Key Location/Qualifiers  
 FT CDS 1..3432  
 FT /tag= a  
 FT /product= "splice variant of MATER"  
 EN EPI285964-A2.  
 PD 26-FEB-2003.  
 PD 12-JUL-2002; 2002EP-00090246.  
 PR 10-AUG-2001; 2001DE-01039874.  
 PA (SCHD ) SCHERING AG.  
 PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;  
 DR WP1; 2003-302814/30.  
 DR P-PSDB; ADA45221.  
 XX New nucleic acid for human maternal antigen that embryos require protein,  
 PT useful e.g. in diagnosis and treatment of female infertility.  
 PS Claim 1; Page 17-18; 31pp; German.  
 CC This invention describes a novel human MATER (maternal antigen that  
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase  
 CC associated with apoptosis, a defect in MATER activity causes growth  
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the  
 CC MATER gene is found predominantly in ovary, testis and placenta. The  
 CC products of the invention have antiinfertility, contraceptive,  
 CC antiinflammatory, immunosuppressive and gynaecological activity.  
 CC Effectors of the MATER polypeptide are used to treat infertility.  
 CC associated with endometriosis and also ovarian dysfunction, autoimmune  
 CC premature ovarian failure, inflammation, autoimmune diseases and female  
 CC infertility, and as contraceptives. Measurements of autoantibodies  
 CC against MATER in body samples or MATER protein or mRNA in eggs is useful  
 CC for diagnosis of female infertility, particularly by detecting mutations  
 CC in the MATER gene with a DNA chip. This sequence represents a splice  
 CC variant of human MATER in which exon 4 of the usual form is absent.  
 XX  
 SQ Sequence 3830 BP; 943 A; 946 C; 1041 G; 900 T; 0 U; 0 Other;

Query Match 77.8%; Score 3139; DB 8; Length 3830;  
 Best Local Similarity 96.0%; Pred. No. 0;  
 Matches 3282; Conservative 0; Mismatches 30; Indels 105; Gaps 2;

QY 1 ATGGAAGAGACAATATGCTTCCAGTACGAGGCTGCAATGTCTCTATGAG 60  
 DB 1 ATGGAAGAGACAATATGCTTCCAGTACGAGGCTGCAATGTCTCTATGAG 60  
 QY 61 CTAGACAAGAGAAATTCAGACATTCAGAGATTACTTAAGAAGAAATCTTCGAATCG 120  
 DB 61 CTAGACAAGAGAAATTCAGACATTCAGAGATTACTTAAGAAGAAATCTTCGAATCG 120  
 QY 121 ACCAGATGCTATTTCCACAGTTTGAAGTGAAGATGCCAAGCGGATGCTGGCACTC 180  
 DB 121 ACCAGATGCTATTTCCACAGTTTGAAGTGAAGATGCCAAGCGGATGCTGGCACTC 180  
 QY 181 CTTTGATGAGATTAATGAGCATGCTGGCTGAGCTACGTCATTAAGATCTTTGAA 240  
 DB 181 CTTTGATGAGATTAATGAGCATGCTGGCTGAGCTACGTCATTAAGATCTTTGAA 240  
 QY 241 AACATGAACCTGGCAACCTCTCGAGAGAGCGAGATGATGAAAAAATTCACAA 300  
 DB 241 AACATGAACCTGGCAACCTCTCGAGAGAGCGAGATGATGAAAAAATTCACAGAA 300

OY		301	GCT-----ATGGAAAGAAGTGGCAACGCGCAGACAAGAACAC-	344
Db		301	GATCTGTGAAGCAACGATGACTGCACAGACCAGCAAAGAAAAGTGCCAGAAAAATAA	360
OY		345	-----	344
Db		361	TATGGCATGACTTAACCTTACTTTGGGGGTGTGCATCTCGACTCCGAATTATTAACAC	420
OY		345	-----AGAAATTTCAACAAGCTATGGACAAGAGGT	375
Db		421	AAGTAGTTGGAAATTCATTTCTTTTGGCAAAATTTCCAAGCTATGGACAAGAGGT	480
OY		376	GCCACAGAGACGACACAGAAACAAGGACATGGAGGTGACATNTGGGACTACAAAGGT	435
Db		481	GCCACAGACGACAGACACAGAAACAAGGACATGGAGGTGACATNTGGGACTACAAAGGT	540
OY		436	CACGTGATGACCAAAATTCGCTGAGGAGGAGATGTACGTCTGATGTTTTGAAAACACTGT	495
Db		541	CACGTGATGACCAAAATTCGCTGAGGAGGAGATGTACGTCTGATGTTTTGAAAACACTGT	600
OY		496	GCTGACTGACCCGGAATGCAAAAGTTGGCTGTGCTTTTGATTACAGCCGGTGGGCTTC	555
Db		601	GCTGACTGACCCGGAATGCAAAAGTTGGCTGTGCTTTTGATTACAGCCGGTGGGCTTC	660
OY		556	CGGCTCCGCACGGTGTCTCTGACCGGAAAGTACAGAAATTTGGGAAATTCGGCTTACCCAGA	615
Db		661	CGGCTCCGCACGGTGTCTCTGACCGGAAAGTACAGAAATTTGGGAAATTCGGCTTACCCAGA	720
OY		616	AGGATCGATGCTGTGCTGGGCGCAAGTGGACTTACCAAGGAAATGTTCTCTCAAGCTTC	675
Db		721	AGGATCGATGCTGTGCTGGGCGCAAGTGGACTTACCAAGGAAATGTTCTCTCAAGCTTC	780
OY		676	TTCTCCCCCTTGAAGAGATGACGCGGAAGAAGAGACAGTGTCAACAGATTCACTCC	735
Db		781	TTCTCCCCCTTGAAGAGATGACGCGGAAGAAGAGAGAGATGTCAACAGATTCACTCC	840
OY		736	AGGAGTGGCCAGACTCCAGGCTCCGGTACGACGAGATCATGTCCGACCAAGAAAGCTG	795
Db		841	AGGAGTGGCCAGACTCCAGGCTCCGGTACGAGAGATCATGTCCGACCAAGAAAGCTG	900
OY		796	TTGTTTATCATTTAGAGGTGGATGACCGGGCTCTGTCTTCAACAATGACCAAAAGCTC	855
Db		901	TTGTTTATCATTTAGAGGTGGATGACCGGGCTCTGTCTTCAACAACAATGACCAAAAGCTC	960
OY		856	TGCAAAAGACTGGGCTGAGAAGCAGCTCCGTTCAACCTCATACGAGCTGTGCTGAGAG	915
Db		961	TGCAAAAGACTGGGCTGAGAAGCAGCTCCGTTCAACCTCATACGAGCTGTGCTGAGAG	1020
OY		916	GTCCTGTCTCCCTGAGTCTCTCTCTGATGTCAACCGTCAGAGAGCTGGGCAACAGAAAGCTC	975
Db		1021	GTCCTGTCTCCCTGAGTCTCTCTCTGATGTCAACCGTCAGAGAGCTGGGCAACAGAAAGCTC	1080
OY		976	AAGTCAGAGGTGTGTCTCTCCCGTTAACTGTTAAGTAAAGAAATTCCTCGGGGAAACAAGA	1035
Db		1081	AAGTCAGAGGTGTGTCTCTCCCGTTAACTGTTAAGTAAAGAAATTCCTCGGGGAAACAAGA	1140
OY		1036	ATCCACTTGTCTCTTGAGCGCGGATTTGGTGAACATCAGAAAGACACAAGGTTGGTGTG	1095
Db		1141	ATCCACTTGTCTCTTGAGCGCGGATTTGGTGAACATCAGAAAGACACAAGGTTGGTGTG	1200
OY		1096	ATCATGAACAAACCGTAGCTGTGCACAGATGCCAGATGCCCGCCGTCGTGGGCTCTCTCATC	1155
Db		1201	ATCATGAACAAACCGTAGCTGTGCACAGATGCCAGATGCCCGCCGTCGTGGGCTCTCTCATC	1260
OY		1156	TGCGTGGCCCTTGACGCTGACAGACGTGTGGGGGAGAGCTGTGCCCTTTCAACCAACG	1215
Db		1261	TGCGTGGCCCTTGACGCTGACAGACGTGTGGGGGAGAGCTGTGCCCTTTCAACCAACG	1320
OY		1216	CTCACAGGCTGACGCGGCTTTTGTTTCAATCAAGCTCAACCCCTCGAAGGCGTGTCCGG	1275
Db		1321	CTCACAGGCTGACGCGGCTTTTGTTTCAATCAAGCTCAACCCCTCGAAGGCGTGTCCGG	1380
OY		1276	CGGTGTCTCAATCTTGAGAAAGATTTGTCTCTGAAGCGCTTTCGCGGATATGGCTGTGAG	1335

Db	1381	CGCTGTCTCAATCTGAGAGAAAAGATTCTCTGAAGCGCTTCTGCGGTATGGCTGTGGAG	14440
Qy	1336	GGAGTGTGAAATAGGAAGTCAGTGTTTGATGTGACGACCTCATGGTTCAAGACTCGAG	13995
Db	1441	GGAGTGTGAAATAGGAAGTCAGTGTTTGATGTGACGAGTCAATGGTTCAAGAACTCGAG	15000
Qy	1386	GAGTCTGAGCTCCGTGCTCTGTTTCACTGAAACTCTTCTTCCAGACAGCTATGTGAG	14557
Db	1501	GAGTCTGAGCTCCGTGCTCTGTTTCACTGAAACTCTTCTTCCAGACAGCTACTGTGAG	15660
Qy	1456	GAGTACTACACTTCTTCCACTGAGTCCAGGACTCTGTGCGCGCTGTACTAGCTG	15151
Db	1561	GAGTACTACACTTCTTCCACTGAGTCCAGGACTCTGTGCGCGCTGTACTAGCTG	16202
Qy	1516	TTAGAAGGCGCTGGAATCGAATCGACCGACTCTGCGCTCTGTAAGTTGAAGACAAAGAG	15757
Db	1621	TTAGAAGGCGCTGGAATCGAATCGACCGACTCTGCGCGCTCTGTAAGTTGAAGACAAAGAG	16818
Qy	1576	TCATAGAGCTTTAAACAGGACGCTTCCATATCCACTGCTTTGGATGAAGCTTCTTG	16337
Db	1681	TCATAGAGCTTTAAACAGGACGCTTCCATATCCACTGCTTTGGATGAAGCTTCTTG	17440
Qy	1636	TTTGCGCTCTGAGAGGAGAAAGTAAAGGAGGCTCACTGAGAGTCTGCTGGGCTGTCCGCT	16997
Db	1741	TTTGCGCTCTGAGAGGAGAAAGTAAAGGAGGCTCACTGAGAGTCTGCTGGGCTGTCCGCT	18000
Qy	1696	CCCCGTGGGGTGAAGACAGAACTTCTGCACTGGGCTCTCTGTTGGGTCAACAGCTAAAT	17555
Db	1801	CCCCGTGGGGTGAAGACAGAACTTCTGCACTGGGCTCTCTGTTGGGTCAACAGCTAAAT	18660
Qy	1756	GCCACACCCCGAGAGACACCTCTGAGCGCTTCCACTGTCTTTTGAGACTCAAGACAA	18151
Db	1861	GCCACACCCCGAGAGACACCTCTGAGCGCTTCCACTGTCTTTTGAGACTCAAGACAA	19202
Qy	1816	GAGTTGTTCCTGGCTTGAACAAGCTTCCAAAGATGTGGCTTCCGATTTAACAGAAC	18757
Db	1921	GAGTTGTTCCTGGCTTGAACAAGCTTCCAAAGATGTGGCTTCCGATTTAACAGAAC	19800
Qy	1876	CTGGACTTGATAGCATCTTCCCTCTGCGCTCGAGCACTGCTCCGTATTTGCGGAAAAATTGG	19353
Db	1981	CTGGACTTGATAGCATCTTCCCTCTGCGCTCGAGCACTGCTCCGTATTTGCGGAAAAATTGG	20404
Qy	1936	GTGATGTCAAAAGGATCTTCCCAAGATAGTCCGCTGAAGCATGTCTGTGGTCCCT	19995
Db	2041	GTGATGTCAAAAGGATCTTCCCAAGATAGTCCGCTGAAGCATGTCTGTGGTCCCT	21000
Qy	1996	CTATGATGCGGGATTAAGACCTTCAATTAAGAGCATGTGGAAAGTTTGTCTCCATGCTT	20555
Db	2101	CTATGATGCGGGATTAAGACCTTCAATTAAGAGCATGTGGAAAGTTTGTCTCCATGCTT	21660
Qy	2056	GGACCCCAACCAACCTGCGGCACTGGACCTGCGGACAGCACTCTGACAGAGCGGCGC	21151
Db	2161	GGACCCCAACCAACCTGCGGCACTGGACCTGCGGACAGCACTCTGACAGAGCGGCGC	22202
Qy	2116	ATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATACAGACCTGATGTTT	21757
Db	2221	ATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATACAGACCTGATGTTT	22800
Qy	2176	AGAAATGCAAGATTAACCTCTGTGTGACAGCACTTGTGAGAAATGTCTATGGCAACCT	22357
Db	2281	AGAAATGCAAGATTAACCTCTGTGTGACAGCACTTGTGAGAAATGTCTATGGCAACCT	23408
Qy	2236	AACTTAAGATCCCTCAACTTGGGAGGACCCCACTGAAGGAAGAGATGAAGATGGCG	22959
Db	2341	AACTTAAGATCCCTCAACTTGGGAGGACCCCACTGAAGGAAGAGATGAAGATGGCG	24000
Qy	2296	TGTGAAGCTTTAAACACCCAAATGTTTGTGTGAGTCTTTGAAGCTGATTTGCTGTGGA	23555
Db	2401	TGTGAAGCTTTAAACACCCAAATGTTTGTGTGAGTCTTTGAAGCTGATTTGCTGTGGA	24606
Qy	2356	TTGACCAATGCTGTAACTGGAAGATCTCCCAATCTTTAGACTTCCCCAGCTTGA	24151

Db 2461 TTGACCCATGCTCTTACCTGTAAGATCTCCCAAACTCTTACGACCTCCCGACGCTGAAA 2520  
 Qy 2416 TCTCTGAGCCTGGACGAGAAAGAGTGA CAGACCGAGGAGTAATGCTCTGAGTATGCC 2475  
 Db 2521 TCTCTGAGCCTGGACGAGAAAGAGTGA CAGACCGAGGAGTAATGCTCTGAGTATGCC 2580  
 Qy 2476 TTGAGAGTCTCCCAAGTGGCCCTGCAAGAGTGA TACTGAGAGACTGTGGATCA CAGACC 2535  
 Db 2581 TTGAGAGTCTCCCAAGTGGCCCTGCAAGAGTGA TACTGAGAGACTGTGGATCA CAGACC 2640  
 Qy 2536 AGCGGTTCCGAGAGTCTGGCTCAGCCCTCTG CAGCAACCGGAGCTTGA CACACTGTGC 2595  
 Db 2641 AGCGGTTCCGAGAGTCTGGCTCAGCCCTCTG CAGCAACCGGAGCTTGA CACACTGTGC 2700  
 Qy 2596 CTATCCAA CACAGCTGGGGAGAGAGTGAATCTACTGTGTGCATGATGAGGCTT 2655  
 Db 2701 CTATCCAA CACAGCTGGGGAGAGAGTGAATCTACTGTGTGCATGATGAGGCTT 2760  
 Qy 2656 CCCCACTGAGTCTG CAGAGGCTGATGCTGAATCAGTCCACCTGGA CACGCTGGCTGT 2715  
 Db 2761 CCCCACTGAGTCTG CAGAGGCTGATGCTGAATCAGTCCACCTGGA CACGCTGGCTGT 2820  
 Qy 2716 GGTTCCTTTCGACTTGGCTTATGGGTATCTGATGCTGAGCGACCTGAGCTTACATG 2775  
 Db 2821 GGTTCCTTTCGACTTGGCTTATGGGTATCTGATGCTGAGCGACCTGAGCTTACATG 2880  
 Qy 2776 AACCTGTGGAAGACATAGGCGGTGAAGCTTGTGTGTCGAGGTGATGAGAGAA CCACTTGT 2835  
 Db 2881 AACCTGTGGAAGACATAGGCGGTGAAGCTTGTGTGTCGAGGTGATGAGAGAA CCACTTGT 2940  
 Qy 2836 CATCTCCAGAGACTGGAAGTGTGAATGATGATCTCAGCCGCGCGTGTGAGAGTCTG 2895  
 Db 2941 CATCTCCAGAGACTGGAAGTGTGAATGATGATCTCAGCCGCGCGTGTGAGAGTCTG 3000  
 Qy 2896 TCTCTGATGATCTG CAGAGACGACACTGAAAGCCTGAGATCTGACGAA CAAATGCCCTG 2955  
 Db 3001 TCTCTGATGATCTG CAGAGACGACACTGAAAGCCTGAGATCTGACGAA CAAATGCCCTG 3060  
 Qy 2956 GGTGACGCTGGGGTGTGCTGCGTGTGCGAGGGA CTGAAGCAAAAGAACAGTGTCTGACG 3015  
 Db 3061 GGTGACGCTGGGGTGTGCTGCGTGTGCGAGGGA CTGAAGCAAAAGAACAGTGTCTGACG 3120  
 Qy 3016 AGACTCGGCTGGAAGGCACTGTGACTGACTTCTGATTCCTGTGAGGCACTCTCTTGGCC 3075  
 Db 3121 AGACTCGGCTGGAAGGCACTGTGACTGACTTCTGATTCCTGTGAGGCACTCTCTTGGCC 3180  
 Qy 3076 CTTTCTGCAACCGGCACTGTGACTGACTTCTGATTCCTGTGAGGCACTCTCTTGGCC 3135  
 Db 3181 CTTTCTGCAACCGGCACTGTGACTGACTTCTGATTCCTGTGAGGCACTCTCTTGGCC 3240  
 Qy 3136 GGAATGATGAAGCTGTGTTGGGCTTGGCTGTGCTGCTGCAAGTCTTACAGATAATTGGG 3195  
 Db 3241 GGAATGATGAAGCTGTGTTGGGCTTGGCTGTGCTGCTGCAAGTCTTACAGATAATTGGG 3300  
 Qy 3196 CTGTGAAATGAGCACTACCTGTGCAAA TAAAGAACTGCTGAGAGAA GTGCACTATC 3255  
 Db 3301 CTGTGAAATGAGCACTACCTGTGCAAA TAAAGAACTGCTGAGAGAA GTGCACTATC 3360  
 Qy 3256 AAGCCCCGAGTCTGTAATGAGAGGTAGTGGCA TTTCTTTGATGAAGATGACCGACAC 3312  
 Db 3361 AAGCCCCGAGTCTGTAATGAGAGGTAGTGGCA TTTCTTTGATGAAGATGACCGACAC 3417

RESULT 6  
 AAD41224  
 ID AAD41224 standard; cDNA, 3489 BP.

XX AAD41224;  
 XX AC  
 XX 30-OCT-2002 (first entry)  
 XX DT  
 XX Human EMBRY-1 cDNA.  
 XX

KW Human; embryogenesis associated protein; AIDS; reproductive disorder;  
 KW infertility; endometriosis; endometrial tumour; inflammatory disorder;  
 KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;  
 KW ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;  
 KW EMBRY-1; allergy; gene therapy; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH 1..3489  
 FT /tag= a  
 FT CDS /product= "EMBRY-1 protein"  
 FT  
 FN MO200248362-A2.  
 XX  
 PD 20-JUN-2002.  
 XX  
 PR 14-NOV-2001; 2001MO-US043956.  
 XX  
 PR 15-NOV-2000; 2000US-0249407P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Rankumar J, Arvizu C;  
 P1  
 XX MPI; 2002-537629/57.  
 DR P-PSDB; AAE25053.  
 XX  
 PT New polypeptides of human embryogenesis associated proteins for screening  
 PT modulators useful for treating or preventing disorders e.g.  
 PT endometriosis, infertility, allergy, preeclampsia.  
 XX  
 PS Claim 58; Page 95-96; 97pp; English.

CC The invention relates to human embryogenesis associated proteins (EMBRY)  
 CC and nucleic acid molecules encoding such proteins. EMBRY sequences are  
 CC useful for screening modulators useful for treating or preventing  
 CC disorders associated with abnormal expression of EMBRY. The disorders  
 CC treated include reproductive disorders such as infertility,  
 CC endometriosis, endometrial or ovarian tumour, autoimmune/inflammatory  
 CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,  
 CC contact dermatitis; disorders of the placenta such as preeclampsia,  
 CC abruptio placentae etc. Sequences of the invention are also useful for  
 CC analysing a proteome of a tissue or a cell type. EMBRY proteins are  
 CC useful as immunogens for preparing antibodies. Polynucleotides of the  
 CC invention are useful for creating knockin humanised animals or transgenic  
 CC animals to model human diseases. They are also used in gene therapy. The  
 CC present sequence is human EMBRY-1 cDNA  
 CC  
 XX  
 SQ Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;

Query Match 77.7%; Score 3133.6; DB 6; Length 3489;  
 Best Local Similarity 95.2%; Pred. No. 0;  
 Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

Qy 1 ATGGAAGAGCAAAATGCTACCTTTCCAGTACGGGCTGCAATGCTCTATGAG 60  
 Db 1 ATGGAAGAGCAAAATGCTACCTTTCCAGTACGGGCTGCAATGCTCTATGAG 60  
 Qy 61 CTAGACAAGAGAAATTTCAAGCACTTCAAGAA TTAAGAAATCTTCAGAAATCG 120  
 Db 61 CTAGACAAGAGAAATTTCAAGCACTTCAAGAA TTAAGAAATCTTCAGAAATCG 120  
 Qy 121 ACCGATGCTCTATTTCCAGAGTTTGAATGAGAAATCCAA CCGTGAATGCTGGCACTC 180  
 Db 121 ACCGATGCTCTATTTCCAGAGTTTGAATGAGAAATCCAA CCGTGAATGCTGGCACTC 180  
 Qy 181 CTTTGCATGAGATTTATGAGAGATCGCTGGCTGAGCTAGTCCATTAGCATCTTTGAA 240  
 Db 181 CTTTGCATGAGATTTATGAGAGATCGCTGGCTGAGCTAGTCCATTAGCATCTTTGAA 240  
 Qy 241 AACATGAACCTGCCAACCTCTCGGAGAGGCA CGGGATGACATGAATA----- 289

Db 241 AACATGAACCTGCAACCTCTCGAGAAAGCAGGGATGATGATAAAATTCAACAGAA 300  
Qy 290 ----- 289  
Db 301 GATCCTGAACAGATGA CTGACCAAGGACCAAGCAAGAAAAAGTGCAGAAAAATAA 360  
Qy 290 ----- 289  
Db 361 TATGGCATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAACAC 420  
Qy 290 ----- AAATTTCACAGCTATGAGACAGAGAGT 318  
Db 421 AAGTATGTTGAATTCAATCTTCTTTGAGAAATTTCAAGGCTATGAGAACAGAGAGT 480  
Qy 319 GCCACAGCAGCAGAGACAGAGAACAGAAATTTTCAAGGCTATGAGACAGAGAGTGC 378  
Db 481 GCCACAGCAGCAGAGACAGAGAACAGAAATTTTCAAGGCTATGAGAACAGAGAGTGC 540  
Qy 379 ACAGCAGCAGAGACAGAGAACAGAGACATGAGGTCACATGAGGACTACAGAGTGC 438  
Db 541 ACAGCAGCAGAGACAGAGAACAGAGACATGAGGTCACATGAGGACTACAGAGTGC 600  
Qy 439 GTGATGACCAATTGCTGAGAGAGAGATGATAGTCTGATTTTGAATAACATCTGCT 498  
Db 601 GTGATGACCAATTGCTGAGAGAGAGATGATAGTCTGATTTTGAATAACATCTGCT 660  
Qy 499 GACTGGCCGGAATGCAAACTTTGGCTGCTTTTGAATTCAGACCCGGTGGGCTTCCGG 558  
Db 661 GACTGGCCGGAATGCAAACTTTGGCTGCTTTTGAATTCAGACCCGGTGGGCTTCCGG 720  
Qy 559 CCTGCGAGGTGGTCTGTGACGGAAAGTCAGGAATTTGGGAAATCGGCTTACCCAGAGG 618  
Db 721 CCTGCGAGGTGGTCTGTGACGGAAAGTCAGGAATTTGGGAAATCGGCTTACCCAGAGG 780  
Qy 619 ATCGTGTGTGCTGGGCGCAGAGTGAAGTCTTACAGAGGATTTCTCTACCTCTTCTC 678  
Db 781 ATCGTGTGTGCTGGGCGCAGAGTGAAGTCTTACAGAGGATTTCTCTACCTCTTCTC 840  
Qy 679 CTCCCGTTAGAGAGATGACGCGGAAAGAGAGAGAGATGACAGAGTTCACTCCAGG 738  
Db 841 CTCCCGTTAGAGAGATGACGCGGAAAGAGAGAGAGATGACAGAGTTCACTCCAGG 900  
Qy 739 GAGTGGCCAGATCCCAAGCTCCGGTGAACGAGATCATGTCCCGACCAAGAGGCTGTTG 798  
Db 901 GAGTGGCCAGATCCCAAGCTCCGGTGAACGAGATCATGTCCCGACCAAGAGGCTGTTG 960  
Qy 799 TTCAATCATGACGTTTTCATGATGATGAGCTCTGCTCTTCAACAATGACCAAGGCTTGC 858  
Db 961 TTCAATCATGACGTTTTCATGATGATGAGCTCTGCTCTTCAACAATGACCAAGGCTTGC 1020  
Qy 859 AAAAGCTGGGCTGAGAGAGAGCTTCGTTCACTCATACGAGTCTGTGAGAGAGTGC 918  
Db 1021 AAAAGCTGGGCTGAGAGAGAGCTTCGTTCACTCATACGAGTCTGTGAGAGAGTGC 1080  
Qy 919 CTGCTCCCTGATGATCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 978  
Db 1081 CTGCTCCCTGATGATCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Qy 979 TCAGAGTGTGTGTCTCCCGTTACCTGTTAGTTAGAGAAATCTCCGGGGAAACAAAGAAATC 1038  
Db 1141 TCAGAGTGTGTGTCTCCCGTTACCTGTTAGTTAGAGAAATCTCCGGGGAAACAAAGAAATC 1200  
Qy 1039 CACTTGTCTCTTGAACGCGGGAGTTGTGATGATGATGATGATGATGATGATGATGATGAT 1098  
Db 1201 CACTTGTCTCTTGAACGCGGGAGTTGTGATGATGATGATGATGATGATGATGATGATGAT 1260  
Qy 1099 ATGAACAACCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1158  
Db 1261 ATGAACAACCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Qy 1159 GTGGCCCTGACAGTCAAGAGAGTGTGGGGAGAGAGTGTGGCCCTTCAACAAGAGCTGC 1218  
Db 1321 GTGGCCCTGACAGTCAAGAGAGTGTGGGGAGAGAGTGTGGCCCTTCAACAAGAGCTGC 1380

Qy 1219 AAGGCTGACAGCGCCCTTTTGTGTTTCATGAGCTCAACCCCTCGAGGAGTGTCCGGGC 1278  
Db 1381 AAGGCTGACAGCGCCCTTTTGTGTTTCATGAGCTCAACCCCTCGAGGAGTGTCCGGGC 1440  
Qy 1279 TGTCTCAATCTGAGAGAAAGTGTCTGAGAGGCTTCTGCGTATGCTGTGAGAGGA 1338  
Db 1441 TGTCTCAATCTGAGAGAAAGTGTCTGAGAGGCTTCTGCGTATGCTGTGAGAGGA 1500  
Qy 1339 GTGTGAATGAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398  
Db 1501 GTGTGAATGAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
Qy 1339 TGTGAGCTCCGCTCTGCTTTTCACTGAACATTCCTTCTCCAGAGAGCAGCAGTGAAGAG 1458  
Db 1561 TGTGAGCTCCGCTCTGCTTTTCACTGAACATTCCTTCTCCAGAGAGCAGCAGTGAAGAG 1620  
Qy 1459 TACTACACCTTCTTCCACTCAGTCTCCAGAGCTTCTGCGCGCTTACTACGTTTA 1518  
Db 1621 TACTACACCTTCTTCCACTCAGTCTCCAGAGCTTCTGCGCGCTTACTACGTTTA 1680  
Qy 1519 GAGGCTGGAATGAGAGCAGCTCTGCGCTCTGTAAGTTGAGAGACAAAGAGTCC 1578  
Db 1681 GAGGCTGGAATGAGAGCAGCTCTGCGCTCTGTAAGTTGAGAGACAAAGAGTCC 1740  
Qy 1579 ATGAGCTTAAACAGGAGGCTTCAATCCATGCTGTTGATGATGATGATGATGATGATGAT 1638  
Db 1741 ATGAGCTTAAACAGGAGGCTTCAATCCATGCTGTTGATGATGATGATGATGATGATGAT 1800  
Qy 1639 GAGCTGTGAGAGAGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1698  
Db 1801 GAGCTGTGAGAGAGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860  
Qy 1699 CTGGGGGTGAAGCAGAACTTCTGACATGAGTCTCTGTTGGGTGAGCAGCCTTAATGCC 1758  
Db 1861 CTGGGGGTGAAGCAGAACTTCTGACATGAGTCTCTGTTGGGTGAGCAGCCTTAATGCC 1920  
Qy 1759 ACCACCCAGAGAGACCTCTGAGAGCTTCCATGCTCTTTTGAAGCTCAAGACAAAGAG 1818  
Db 1921 ACCACCCAGAGAGACCTCTGAGAGCTTCCATGCTCTTTTGAAGCTCAAGACAAAGAG 1980  
Qy 1819 TTTGTTGCTTGGGATTTAAACAGCTTCCAAAGAGTGGCTTCCGATTAACAGAACTTG 1878  
Db 1981 TTTGTTGCTTGGGATTTAAACAGCTTCCAAAGAGTGGCTTCCGATTAACAGAACTTG 2040  
Qy 1879 GACTTATGATGATCTTCTCTGCTCCAGACATGCTCCGATTTTGGGAAAAATTCGGGTG 1938  
Db 2041 GACTTATGATGATCTTCTCTGCTCCAGACATGCTCCGATTTTGGGAAAAATTCGGGTG 2100  
Qy 1939 GATGTCAAGGAGATCTTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1998  
Db 2101 GATGTCAAGGAGATCTTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
Qy 1999 TGAATGGGAGTAAAGCCCTCATTTGAGAGAGAGTGGAGAAATTTCTGCTCCATGCTTGGC 2058  
Db 2161 TGAATGGGAGTAAAGCCCTCATTTGAGAGAGAGTGGAGAAATTTCTGCTCCATGCTTGGC 2220  
Qy 2059 ACCACCCAGACCTGCGGAGCCTGAGACCTGAGGAGAGATCTTGAACAGAGCGGGCCATG 2118  
Db 2221 ACCACCCAGACCTGCGGAGCCTGAGACCTGAGGAGAGATCTTGAACAGAGCGGGCCATG 2280  
Qy 2119 AAGACCTGTGTGCCAAGTGAAGTCACTTGAAGATGACAGACCTGATGTTTGA 2178  
Db 2281 AAGACCTGTGTGCCAAGTGAAGTCACTTGAAGATGACAGACCTGATGTTTGA 2340  
Qy 2179 AATGACAGATTAACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2238  
Db 2341 AATGACAGATTAACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
Qy 2239 CTAAGATCCCTCAACTTGGAGAGCAGCAGCTGAGAGAAAGATGATGATGATGATGATGAT 2298  
Db 2401 CTAAGATCCCTCAACTTGGAGAGCAGCAGCTGAGAGAAAGATGATGATGATGATGATGAT 2460



```

QY 2299 GAAAGCTTTAAAAACCCAAAAATGTTTGTGGAGTCTTTGAGGCTGGAATGCTGATG 2358
DB 2461 GAAAGCTTTAAAAACCCAAAAATGTTTGTGGAGTCTTTGAGGCTGGAATGCTGATG 2520
QY 2359 ACCCATGCTTGTACCTGAAAGTCTCCCAATCTCTTACGACCTCCCGACCTGAAATCT 2418
DB 2521 ACCCATGCTTGTACCTGAAAGTCTCCCAATCTCTTACGACCTCCCGACCTGAAATCT 2580
QY 2419 CTGAGCTTGGAGAAAAGAGTGAAGAGCCAGAGGAGTATAGCCCTCTGAGTATGCTTGG 2478
DB 2581 CTGAGCTTGGAGAAAAGAGTGAAGAGCCAGAGGAGTATAGCCCTCTGAGTATGCTTGG 2640
QY 2479 AGAGTCTCCAGATGCGCTCTGAGAGCTGATATGAGAGACTGTGAGCATACAGCCACG 2538
DB 2641 AGAGTCTCCAGATGCGCTCTGAGAGCTGATATGAGAGACTGTGAGCATACAGCCACG 2700
QY 2539 GGTGTCAGAGTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAG 2598
DB 2701 GGTGTCAGAGTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAG 2760
QY 2599 TCCAAACACAGCTGAGGAGAGAGGAGTAAATCTACTGTGTGATGATGATGAGGCTTCC 2658
DB 2761 TCCAAACACAGCTGAGGAGAGAGGAGTAAATCTACTGTGTGATGATGATGAGGCTTCC 2820
QY 2659 CACTGTAGTCTGAGAGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 2718
DB 2821 CACTGTAGTCTGAGAGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 2880
QY 2719 TTTCTTGTGACCTTGTGCTTATGAGTATGATGATGATGATGATGATGATGATGATGATG 2778
DB 2881 TTTCTTGTGACCTTGTGCTTATGAGTATGATGATGATGATGATGATGATGATGATGATG 2940
QY 2779 CCTGTGAAAGCAATGAGGCTGAGCTGATGAGCTGATGAGCTGATGAGCTGATGAGCTGAT 2838
DB 2941 CCTGTGAAAGCAATGAGGCTGAGCTGATGAGCTGATGAGCTGATGAGCTGATGAGCTGAT 3000
QY 2839 CTCCAGACCTGAGATGTTGTTAAAGTGTATCTCAACCGCGGTGCTGTGAGAGTCTGCC 2898
DB 3001 CTCCAGACCTGAGATGTTGTTAAAGTGTATCTCAACCGCGGTGCTGTGAGAGTCTGCC 3060
QY 2899 TGTGTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2958
DB 3061 TGTGTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
QY 2959 GACGCTGAGGAGTGTGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3018
DB 3121 GACGCTGAGGAGTGTGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
QY 3019 CTGCGGTTGAAAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3078
DB 3181 CTGCGGTTGAAAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
QY 3079 TCTGTGACACCGGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3138
DB 3241 TCTGTGACACCGGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
QY 3139 ATGATGAGAGCTGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3198
DB 3301 ATGATGAGAGCTGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
QY 3199 TGAAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3258
DB 3361 TGAAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
QY 3259 CCCCAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
DB 3421 CCCCAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3474

```

RESULT 7  
ADA45218  
ID ADA45218 standard; DNA; 3926 BP.  
XX

```

AC ADA45218;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human MATER DNA.
XX
KW human ; MATER; maternal antigen that embryos require; NTP-ase;
KW nucleoside triphosphate-ase; apoptosis; antifertility; contraceptive;
KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;
KW ovarian dysfunction; autoimmune premature ovarian failure;
KW autoimmune disease; contraceptive; autoantibody; ds; gene.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3489
FT /tag=a
FT /product="MATER"
XX
PN EPI285964-A2.
XX
PD 26-FEB-2003.
XX
PR 12-JUL-2002; 2002EP-00090246.
XX
PR 10-AUG-2001; 2001DE-01039874.
XX
PA (SCHD ) SCHERING AG.
XX
PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
XX
DR WPI; 2003-302814/30.
XX
DR P-PSDB; ADA45219.
XX
PT New nucleic acid for human maternal antigen that embryos require protein,
XX
PS useful e.g. in diagnosis and treatment of female infertility.
XX
PS Claim 1; Page 9-11; 31pp; German.
XX
CC This invention describes a novel human MATER (maternal antigen that
CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
CC associated with apoptosis, a defect in MATER activity causes growth
CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
CC MATER gene is found predominantly in ovary, testis and placenta. The
CC products of the invention have antifertility, contraceptive,
CC antiinflammatory, immunosuppressive and gynaecological activity.
CC Effectors of the MATER polypeptide are used to treat infertility.
CC associated with endometriosis and also ovarian dysfunction, autoimmune
CC premature ovarian failure, inflammation, autoimmune diseases and female
CC infertility, and as contraceptives. Measurements of autoantibodies
CC against MATER in body samples or MATER protein or mRNA in eggs is useful
CC for diagnosis of female infertility, particularly by detecting mutations
CC in the MATER gene with a DNA chip. This sequence encodes the human MATER
CC polypeptide described in the invention.
XX
SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;
XX
Query Match 77.7%; Score 3133.6; DB 8; Length 3926;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;
QY 1 ATGGAAGAGAGCAATGCTCACCTTTCCAGCTACGAGGCTGCAATGATGCTATGAG 60
DB 1 ATGGAAGAGAGCAATGCTCACCTTTCCAGCTACGAGGCTGCAATGATGCTATGAG 60
QY 61 CTAGACAAGAGAAATTTCAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 120
DB 61 CTAGACAAGAGAAATTTCAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 120
QY 121 ACCAGATGCTTATTCAGAGTTGAATGAGAAATGCCAAGTGTGAGATGCTTGCACTC 180
DB 121 ACCAGATGCTTATTCAGAGTTGAATGAGAAATGCCAAGTGTGAGATGCTTGCACTC 180

```



181 CTCTGATGAGTATTTATGAGCATCGCTGGCTGGGCTAGCTCATGATCTTTGA 240  
181 CTCTGATGAGTATTTATGAGCATCGCTGGCTGGGCTAGCTCATGATCTTTGA 240  
241 AACATGAACCTGCGAACCTCTCGAGAGGCGAGGATGATGATGAAA----- 289  
241 AACATGAACCTGCGAACCTCTCGAGAGGCGAGGATGATGATGAAAATTCAAGAA 300  
290 ----- 289  
301 GATCTGAAGCAACGATGATGACCAAGGACCAAGCAAGAAAAGTCCAGAAAATAA 360  
290 ----- 289  
361 TATGCGATGACTAAGCTTATCTTGGGGGTCTGACATCTGACCTCGAATATATAAC 420  
290 ----- 318  
421 AAGTATGTTGAATTCATTCTTTTGAGAAATTTCAAGCTATGGAACAAGAGT 480  
319 GCCACAGCAGCAGACAGAGACAGAAATTTCAAGCTATGGAACAAGAGTGC 378  
481 GCCACAGCAGCAGACAGAGACAGAAATTTCAAGCTATGGAACAAGAGTGC 540  
379 ACAGAGAGAGACAGAAAGACAGAGACATGAGGTGACATGGGACTAGAGAGTAC 438  
541 ACAGAGAGAGACAGAAAGACAGAGACATGAGGTGACATGGGACTAGAGAGTAC 600  
439 GTGATGACCAATTCGCTGAGAGAGAGATGACCTGATTTGAAAAACCTGCT 498  
601 GTGATGACCAATTCGCTGAGAGAGAGATGACCTGATTTGAAAAACCTGCT 660  
499 GACTGGCCGGAATGCAAACTGTCGCTGCTTTGATTCAGACCGGTGGGCTTCGG 558  
661 GACTGGCCGGAATGCAAACTGTCGCTGCTTTGATTCAGACCGGTGGGCTTCGG 720  
559 CCTGCGAGGTGCTCTGACGGAAGTCAAGAAATTCGGAATTCGGCTTCAGAGAG 618  
721 CCTGCGAGGTGCTCTGACGGAAGTCAAGAAATTCGGAATTCGGCTTCAGAGAG 780  
619 ATCGTGTGCTGGGCGCAAGTGAAGTCTTACAGAGGATGTTCTCTCACTCTTC 678  
781 ATCGTGTGCTGGGCGCAAGTGAAGTCTTACAGAGGATGTTCTCTCACTCTTC 840  
679 CTCCCGTTAGAGATGACGCGGAAGAGAGAGAGTCAAGAGTCACTTCAGG 738  
841 CTCCCGTTAGAGATGACGCGGAAGAGAGAGAGTCAAGAGTCACTTCAGG 900  
739 GAGTGGCCAGATCCCAAGCTCCGGTGAACGAGATCATGTCCGACCAAGAAAGCTT 798  
901 GAGTGGCCAGATCCCAAGCTCCGGTGAACGAGATCATGTCCGACCAAGAAAGCTT 960  
799 TTTCATCTGACGCTTTCATGATGATGATGATGATGATGATGATGATGATGATG 858  
961 TTTCATCTGACGCTTTCATGATGATGATGATGATGATGATGATGATGATGATG 1020  
859 AAAAGCTGGGCTGAGAGAGAGCTTCGTTACACCTCATACGAGTCTGCTGAGAG 918  
1021 AAAAGCTGGGCTGAGAGAGAGCTTCGTTACACCTCATACGAGTCTGCTGAGAG 1080  
919 CTGCTCCCTGAGTCTTCTCTGATGATGATGATGATGATGATGATGATGATGATG 978  
1081 CTGCTCCCTGAGTCTTCTCTGATGATGATGATGATGATGATGATGATGATGATG 1140  
979 TCAGAGTGTGATCTCCCGTTACCTGTTAGTTAGAGGATCTCCGGGGAACAAAGAT 1038  
1141 TCAGAGTGTGATCTCCCGTTACCTGTTAGTTAGAGGATCTCCGGGGAACAAAGAT 1200  
1039 CACTTGTCTTTGAGAGCGGAGTTGATGATGATGATGATGATGATGATGATGATG 1098  
1201 CACTTGTCTTTGAGAGCGGAGTTGATGATGATGATGATGATGATGATGATGATG 1260  
1099 ATGAACAACCGTGAAGTCTGACCAAGTCCAGAGTCCCGCGTGGGCTCTCTCATCTGC 1158

1261 ATGAACAACCGTGAAGTCTGACCAAGTCCAGAGTCCCGCGTGGGCTCTCTCATCTGC 1320  
1159 GTGGCTCTGAGCTGACAGACGCTGTGGGGAAGAGCGTGCCTCTCAACCAACGCTC 1218  
1321 GTGGCTCTGAGCTGACAGACGCTGTGGGGAAGAGCGTGCCTCTCAACCAACGCTC 1380  
1219 ACAAGCTGCAACCGCTTTTGTGTTCATGAGCTCAACCTCTGAGGCTGTGTCCGGGC 1278  
1381 ACAAGCTGCAACCGCTTTTGTGTTCATGAGCTCAACCTCTGAGGCTGTGTCCGGGC 1440  
1279 TGTCTCAATCTGAGAGAAAGTGTCTGAGAGCGCTCTGCGGATGAGTGTGAGAGGA 1338  
1441 TGTCTCAATCTGAGAGAAAGTGTCTGAGAGCGCTCTGCGGATGAGTGTGAGAGGA 1500  
1339 GTTGAATTAAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1398  
1501 GTTGAATTAAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1560  
1399 TCTGAGCTCCGTGCTCTGTTTCATGATGATGATGATGATGATGATGATGATGATG 1458  
1561 TCTGAGCTCCGTGCTCTGTTTCATGATGATGATGATGATGATGATGATGATGATG 1620  
1459 TACTACACCTTCTTCAACCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1518  
1621 TACTACACCTTCTTCAACCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1680  
1519 GAGGCTCTGAGAAATGAGACGAGTCTCTGCTCTGATGATGATGATGATGATGATG 1578  
1681 GAGGCTCTGAGAAATGAGACGAGTCTCTGCTCTGATGATGATGATGATGATGATG 1740  
1579 ATGAGCTTAAACAGGACGCTTCAATCACTGCTGTTGATGATGATGATGATGATG 1638  
1741 ATGAGCTTAAACAGGACGCTTCAATCACTGCTGTTGATGATGATGATGATGATG 1800  
1639 GAGCTCTGAGAGAGAGAGTGAAGAGGACATGAGAGTCTGCTGAGGCTGCTCCGTTCC 1698  
1801 GAGCTCTGAGAGAGAGAGTGAAGAGGACATGAGAGTCTGCTGAGGCTGCTCCGTTCC 1860  
1699 CTGGGGGTGAAGCAGAACTTCTGCACTGGGCTCTCTGTTGGGTCAGACGCTTAATGCC 1758  
1861 CTGGGGGTGAAGCAGAACTTCTGCACTGGGCTCTCTGTTGGGTCAGACGCTTAATGCC 1920  
1759 ACCACCCAGAGAGACCTCTGAGAGCTTCACTGCTCTTTCGAGACTCAAGACAAAGAG 1818  
1921 ACCACCCAGAGAGACCTCTGAGAGCTTCACTGCTCTTTCGAGACTCAAGACAAAGAG 1980  
1819 TTTGTTGCTGGGATTAACAGCTTCAAGAGTGGGCTTCCGATTAACAGAGACCTG 1878  
1981 TTTGTTGCTGGGATTAACAGCTTCAAGAGTGGGCTTCCGATTAACAGAGACCTG 2040  
1879 GACTTGAATGATCTTCTCTGCTCTGCTCAAGAGTCCGATTTGCGGAAATTCGAGT 1938  
2041 GACTTGAATGATCTTCTCTGCTCTGCTCAAGAGTCCGATTTGCGGAAATTCGAGT 2100  
1939 GATGTCAAGAGGATCTTCCCAAGAGATGAGTCCGCTGAGGATGCTGTGCTCTCTA 1998  
2101 GATGTCAAGAGGATCTTCCCAAGAGATGAGTCCGCTGAGGATGCTGTGCTCTCTA 2160  
1999 TGAATGGGGAATGAAGCTTCAATGAGAGAGAGGGAATTTCTGCTCAATGCTGGC 2058  
2161 TGAATGGGGAATGAAGCTTCAATGAGAGAGAGGGAATTTCTGCTCAATGCTGGC 2220  
2059 ACCACCCAGAGAGCTGAGAGCTGAGAGCTGAGAGAGCTGAGAGAGAGAGAGAGAG 2118  
2221 ACCACCCAGAGAGCTGAGAGCTGAGAGCTGAGAGAGCTGAGAGAGAGAGAGAGAG 2280  
2119 AAGACCTGTGTGCAAGCTGAGAGCTTCCCAAGATGATGATGATGATGATGATGATG 2178  
2281 AAGACCTGTGTGCAAGCTGAGAGCTTCCCAAGATGATGATGATGATGATGATGATG 2340  
2179 AATGACAGATTAACCTCTGCTGCTGAGAGAGAGTCCGATTAACAGAGAGAGAGAG 2238

```
Db 2341 AATGACAGATTATCCCTGTGTGACGACCTGTGAGAAATGTCATGCGCAACCGTAAC 2400
Qy 2239 CTAAGATCCCTCAACTTGGAGGACCCGACCTGAAGAGATGTAAGATGGCGTGT 2298
Db 2401 CTAAATCCCTCAACTTGGAGGACCCGACCTGAAGAGATGTAAGATGGCGTGT 2460
Qy 2299 GAAGCCTTAAACACCCAAAATGTTTGTGAGTCTTGTGAGGCTGATGCTGTGATG 2358
Db 2461 GAAGCCTTAAACACCCAAAATGTTTGTGAGTCTTGTGAGGCTGATGCTGTGATG 2520
Qy 2359 ACCCATGCTGTACTTGAAGTCTCCCAATCCTTACGACCTCCCGACCTGAAATCT 2418
Db 2521 ACCCATGCTGTACTTGAAGTCTCCCAATCCTTACGACCTCCCGACCTGAAATCT 2580
Qy 2419 CTGAGCCTGAGAGAAACAGAGTGAAGACAGAGGATGATAGCCCTGAGATGAGCTTG 2478
Db 2581 CTGAGCCTGAGAGAAACAGAGTGAAGACAGAGGATGATAGCCCTGAGATGAGCTTG 2640
Qy 2479 AGAGTCTCCAGTGGCCCTGCGAGAGCTGATCTGAGAGACTGTGGCATGACAGCCAG 2538
Db 2641 AGAGTCTCCAGTGGCCCTGCGAGAGCTGATCTGAGAGACTGTGGCATGACAGCCAG 2700
Qy 2539 GGTGGCCAGAGTCTGGGCTCAGCCCTGCTGAGCAACCGGAGCTTGAACAACCTTGCTTA 2598
Db 2701 GGTGGCCAGAGTCTGGGCTCAGCCCTGCTGAGCAACCGGAGCTTGAACAACCTTGCTTA 2760
Qy 2599 TCCAACACAGAGCTGGGAGAGAGAGTGAATCTACTGTGTGATTCATGAGGCTTGCC 2658
Db 2761 TCCAACACAGAGCTGGGAGAGAGAGTGAATCTACTGTGTGATTCATGAGGCTTGCC 2820
Qy 2659 CACTGTACTGTGACAGAGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 2718
Db 2821 CACTGTACTGTGACAGAGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 2880
Qy 2719 TTTCTGTGACTTGGGCTTATGGGTAACTCATGAGGCTGAGGACCTGAGGCTTATGATGATG 2778
Db 2881 TTTCTGTGACTTGGGCTTATGGGTAACTCATGAGGCTGAGGACCTGAGGCTTATGATGATG 2940
Qy 2779 CCTGTGAAGACAAATGAGGCTGATGCTGTGCGAGGCTCATGAGAGAGACAAATCTGTGAT 2838
Db 2941 CCTGTGAAGACAAATGAGGCTTATGCTGTGCGAGGCTCATGAGAGAGACAAATCTGTGAT 3000
Qy 2839 CTCAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2898
Db 3001 CTCAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3060
Qy 2899 TGTGTGATCTGAGAGAGACAACTTGAAGAGGCTGATGATGATGATGATGATGATGATGATG 2958
Db 3061 TGTGTGATCTGAGAGAGACAACTTGAAGAGGCTGATGATGATGATGATGATGATGATGATG 3120
Qy 2959 GACGATGAGGCTGTGCTGCGCTGTGCGAGAGGACTGAAGAGAGAGAGAGAGAGAGAGAGAG 3018
Db 3121 GACGATGAGGCTGTGCTGCGCTGTGCGAGAGGACTGAAGAGAGAGAGAGAGAGAGAGAGAG 3180
Qy 3019 CTCGGGTTGAAGGACTGTGACTGATCTGTGATGCTGTGAGGACTCTCTCTGGCCCTT 3078
Db 3181 CTCGGGTTGAAGGACTGTGACTGATCTGTGATGCTGTGAGGACTCTCTCTGGCCCTT 3240
Qy 3079 TCTGTGCAACCGGCACTTCAACCACTGATGATGATGATGATGATGATGATGATGATGATG 3138
Db 3241 TCTGTGCAACCGGCACTTCAACCACTGATGATGATGATGATGATGATGATGATGATGATG 3300
Qy 3139 ATGATGAAGCTGTGTGGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3198
Db 3301 ATGATGAAGCTGTGTGGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3360
Qy 3199 TGAAGATGAGAGTACCTGTGCAAAATGAAGAGCTGTGAGAGAGTGCAGTACTCAAG 3258
Db 3361 TGAAGATGAGAGTACCTGTGCAAAATGAAGAGCTGTGAGAGAGTGCAGTACTCAAG 3420
Qy 3259 CCCCAGTCTGTATGATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3312
Db 3421 CCCCAGTCTGTATGATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3474
```

```
RESULT 8
ID AAL47135 standard; DNA; 5859 BP.
XX
AC AAL47135;
XX
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP8/Py12 coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virocid;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
EN MO200240668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-EP012545.
XX
PR 15-NOV-2000; 2000DE-01056687.
PR 30-NOV-2000; 2000DE-01059595.
XX
PA (APOT-) APOTEC RES & DEV LTD.
XX
PI Teschopp J, Martinson F;
XX
DR WPI: 2002-427093/45.
DR P-PSDB; MA017863.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
PT proteins and antibodies.
XX
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;
XX
Query Match 75.8%; Score 3057.8; DB 6; Length 5859;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;
Qy 1 ATGAGAGAGAGCAAAATGCTCACCTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 60
Db 1975 ATGAGAGAGAGCAAAATGCTCACCTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 2034
Qy 61 CTAGACAAAGAGAAATTTGAGACTTCAAGAGATTTACTAAAGAGAAATCTTCAAGATCG 120
Db 2035 CTAGACAAAGAGAAATTTGAGACTTCAAGAGATTTACTAAAGAGAAATCTTCAAGATCG 2094
Qy 121 ACCACATGCTCTATTTCCACAGTTTGAATGAGAAATGCAACGAGAGATGCTGGCACTC 180
Db 2095 ACCACATGCTCTATTTCCACAGTTTGAATGAGAAATGCAACGAGAGATGCTGGCACTC 2154
Qy 181 CTCTTGATGATGATTTATGAGATCGCTGCGCTGCGCTAGTCATTAAGATCTTTGAA 240
```

```

Db 2155 CTTTGCAATGAGTATTATGAGCATGCGTGGCTGGGCTAGCTCATTTAGATCTTTGAA 2214
Qy 241 AACATGAACCTGCGAAACCTCTTCGAGAAAGGCACGGATGATGATGAAAA----- 289
Db 2215 AACATGAACCTGCGAAACCTCTTCGAGAAAGGCACGGATGATGATGAAAAAGATTCACCA 2274
Qy 290 ----- 289
Db 2275 GAAGATCCTGAAGCAACGATGATGACCAAGGACCAAGCAAGAAAAAGTCCAGGAATT 2334
Qy 290 -----AAATTCA 297
Db 2335 TCACAAGCTGTGCAACAAGTAGTGCAACGCTGCAAGACAAAGAAACAAGAAATTTTCA 2394
Qy 298 CAAAGTATGAAACAAGAGTGCACACAGACAGACAGAAAGAAACAAGAAATTTTCA 357
Db 2395 CAAAGTATGAAACAAGAGTGCACACAGACAGACAGAAAGAAACAAGAAATTTTCA 2454
Qy 358 GGTATGAAACAAGAGTGCACACAGACAGACAGAAAGAAACAAGACATGAGAGTGAC 417
Db 2455 GGTATGAAACAAGAGTGCACACAGACAGACAGAAAGAAACAAGACATGAGAGTGAC 2514
Qy 418 ACATGGAGCTACAAAGATCAAGTATGACCAAAATTCGCTGAGAGAGAGATGATGCTGT 477
Db 2515 ACATGGAGCTACAAAGATCAAGTATGACCAAAATTCGCTGAGAGAGAGATGATGCTGT 2574
Qy 478 AGTTTGAACCACTGCTGTGATGAGCGGAAATGCAACGTTGGCTGTGCTTTTGT 537
Db 2575 AGTTTGAACCACTGCTGTGATGAGCGGAAATGCAACGTTGGCTGTGCTTTTGT 2634
Qy 538 TCAGACCGGTGGGGCTTCCGGCTGCGACAGGTGTTTTCACAGGAAATTTGGG 597
Db 2635 TCAGACCGGTGGGGCTTCCGGCTGCGACAGGTGTTTTCACAGGAAATTTGGG 2694
Qy 598 AAATGGCTCTACGACGAGAAAGATCGTGTGCTGGGCGCAAGTGTGACTTACAGAGGA 657
Db 2695 AAATGGCTCTACGACGAGAAAGATCGTGTGCTGGGCGCAAGTGTGACTTACAGAGGA 2754
Qy 658 ATGTTCTCTTAAGTCTTCTTCTCCCGTTAGAGAGATGACGCGAAGAGAGAGAGT 717
Db 2755 ATGTTCTCTTAAGTCTTCTTCTCCCGTTAGAGAGATGACGCGAAGAGAGAGAGT 2814
Qy 718 GTCAAGAGTTCATCTCCAGGAGTGGGCGACACTCCAGGCTCCGCTGACGAGATCATG 777
Db 2815 GTCAAGAGTTCATCTCCAGGAGTGGGCGACACTCCAGGCTCCGCTGACGAGATCATG 2874
Qy 778 TCCCGACGAAAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGTCCTC 837
Db 2875 TCCCGACGAAAGGCTGTGTTTCATCATTTGACGGTTTCGATGACCTGGGCTCTGTCCTC 2934
Qy 838 AACATGACACAAAGCTCTGCAAGAAGTGGGCTGAGAAAGCAGCTCCGTTCACTCCATA 897
Db 2935 AACATGACACAAAGCTCTGCAAGAAGTGGGCTGAGAAAGCAGCTCCGTTCACTCCATA 2994
Qy 898 CGCAGTCTGTGAGAGAGTCTGCTCCCTGAGTCTTCTCTGATGTCAAGCTCAGAGAC 957
Db 2995 CGCAGTCTGTGAGAGAGTCTGCTCCCTGAGTCTTCTCTGATGTCAAGCTCAGAGAC 3054
Qy 958 GTGGGACAGAGAACTCAAGTCAAGGTCGATGCTCCCGTTTACCTGTTAGTGAAGA 1017
Db 3055 GTGGGACAGAGAACTCAAGTCAAGGTCGATGCTCCCGTTTACCTGTTAGTGAAGA 3114
Qy 1018 ATCTTCGGGGAACAAGATTCATCTGCTCTTGAAGCGCGGATTTGTAGCATGAGAAG 1077
Db 3115 ATCTTCGGGGAACAAGATTCATCTGCTCTTGAAGCGCGGATTTGTAGCATGAGAAG 3174
Qy 1078 ACACAAGGTTGCTGCTGATCATGAAACAAGTGAAGTGTGACCAAGTCCAGGTTGCC 1137
Db 3175 ACACAAGGTTGCTGCTGATCATGAAACAAGTGAAGTGTGACCAAGTCCAGGTTGCC 3234
Qy 1138 GCCGGGGCTCTCTCATCTGCGGTGGCCCTGACGCTGACGAGACGTGGGGGAGAGCCTC 1197

```

```

Db 3235 GCCGGGGCTCTCTCATCTGCGGTGGCCCTGACGCTGACAGACGTGGGGGAGAGCCTC 3294
Qy 1198 GCCCTTTCAACCAAGCTTCAGAGGCTTGACAGCCCGCTTTTGTGTTTCACTGACCC 1257
Db 3295 GCCCTTTCAACCAAGCTTCAGAGGCTTGACAGCCCGCTTTTGTGTTTCACTGACCC 3354
Qy 1258 CCGAGGCGTGTGCTCCGCGCTGTCAATCTGAGAGAAAGATGTTCTGAGAGCGCTTC 1317
Db 3355 CCGAGGCGTGTGCTCCGCGCTGTCAATCTGAGAGAAAGTGTCTGAGAGCGCTTC 3414
Qy 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAAGAGTCAAGTGTGTTAGTGTGAGACCTC 1377
Db 3415 TGCCGTATGCTGTGAGAGAGTGTGAATAAGAGTCAAGTGTGTTAGTGTGAGACCTC 3474
Qy 1378 ATGTTCAAGGACTCGGGGAGTCTGAGTCCGTGCTGTGTTCACTGAACAATCTTCTC 1437
Db 3475 ATGTTCAAGGACTCGGGGAGTCTGAGTCCGTGCTGTGTTCACTGAACAATCTTCTC 3534
Qy 1438 CAGACAGGCACTGAGAGGATGATCAACCTTTCACCTCAGTCTCAGAGCTTCTGT 1497
Db 3535 CAGACAGGCACTGAGAGGATGATCAACCTTTCACCTCAGTCTCAGAGCTTCTGT 3594
Qy 1498 GCCGCTTTGATCTACGTGTTAGAGGCTGGAAATCGAGCAGGCTCTGCTGTGAC 1557
Db 3595 GCCGCTTTGATCTACGTGTTAGAGGCTGGAAATCGAGCAGGCTCTGCTGTGAC 3654
Qy 1558 GTTGAAGAACAAAGAGTCCATGAGCTTAAACAGCAGGCTTCCATATCACTGCTT 1617
Db 3655 GTTGAAGAACAAAGAGTCCATGAGCTTAAACAGCAGGCTTCCATATCACTGCTT 3714
Qy 1618 TGGANTGAAGCGTTTCTGTTTGGCTGTGAGCGCAAGACGTAAGAGGCCACTGAGGTC 1677
Db 3715 TGGANTGAAGCGTTTCTGTTTGGCTGTGAGCGCAAGACGTAAGAGGCCACTGAGGTC 3774
Qy 1678 CTGCTGGGCTGTCCCGTTCCCTGAGGAGTGAAGAGAGTCTTGCACGAGGATCTCTCG 1737
Db 3775 CTGCTGGGCTGTCCCGTTCCCTGAGGAGTGAAGAGAGTCTTGCACGAGGATCTCTCG 3834
Qy 1738 TTGGGTGACAGCCTTAATGCAACCCACAGAGACACCTGACGCTTCACTGTCTT 1797
Db 3835 TTGGGTGACAGCCTTAATGCAACCCACAGAGACACCTGACGCTTCACTGTCTT 3854
Qy 1798 TTGAGAGTCAAGACAAAGATTTGTTGCTGTGCACTTAAACACTTCCAAAGATGG 1857
Db 3895 TTGAGAGTCAAGACAAAGATTTGTTGCTGTGCACTTAAACACTTCCAAAGATGG 3954
Qy 1858 CTTCCGATTAACCAAGACCTGAGCTTGAATGACATCTTCTCTGCTCCAGACCTGTCCG 1917
Db 3955 CTTCCGATTAACCAAGACCTGAGCTTGAATGACATCTTCTCTGCTCCAGACCTGTCCG 4014
Qy 1918 TATTTGGGAAATTTGGGATGATGTCAAGAGATCTTCCAGAGATGATGCTGCTGAG 1977
Db 4015 TATTTGGGAAATTTGGGATGATGTCAAGAGATCTTCCAGAGATGATGCTGCTGAG 4074
Qy 1978 GCATGTCTGTGTCCTCTTATGAGATGCGGATTAAGACCTCATTTGAGAGACATGGGAA 2037
Db 4075 GCATGTCTGTGTCCTCTTATGAGATGCGGATTAAGACCTCATTTGAGAGACATGGGAA 4134
Qy 2038 GATTTGCTCATGAGCTTGGACCAACCACTGAGGAGCTGAGACCTGGGAGAGAGC 2097
Db 4135 GATTTGCTCATGAGCTTGGACCAACCACTGAGGAGCTGAGACCTGGGAGAGAGC 4194
Qy 2098 ATCTGACAGAGCGGACATGAGAACCTGTGTGCAAGCTGAGGACTCCACCTGCAAG 2157
Db 4195 ATCTGACAGAGCGGACATGAGAACCTGTGTGCAAGCTGAGGACTCCACCTGCAAG 4254
Qy 2158 ATAAGAGCCTGAGTGTAAAGATGACAGATTAACCTCTGTGTGACAGCACTTGGAGA 2217
Db 4255 ATAAGAGCCTGAGTGTAAAGATGACAGATTAACCTCTGTGTGACAGCACTTGGAGA 4314
Qy 2218 ATGCTCATGAGCAACCGGTAACCTTAAGATCCCTCAACTTGGAGGACCAACCTGAAGAA 2277
Db 4315 ATGCTCATGAGCAACCGGTAACCTTAAGATCCCTCAACTTGGAGGACCAACCTGAAGAA 4374

```

```
OY 2278 GAGGATGTAAGATGGCGTGTGTAAGCCTTAAACACCCAAATGTTTGTGGAGTCTTTC 2337
DB 4375 GAGGATGTAAGATGGCGTGTGTAAGCCTTAAACACCCAAATGTTTGTGGAGTCTTTC 4434
OY 2338 AGGCTGATTTGTGTGATTTGAACCCATGCTGTGTAACCTGAAGATCTCCAAATCTTACG 2397
DB 4435 AGGCTGATTTGTGTGATTTGAACCCATGCTGTGTAACCTGAAGATCTCCAAATCTTACG 4494
OY 2398 ACCCTCCAGCGCTGAATTTCTGTAGCCTGTGAGGAAACAAGTGAACAGACAGGAGATA 2457
DB 4495 ACCCTCCAGCGCTGAATTTCTGTAGCCTGTGAGGAAACAAGTGAACAGACAGGAGATA 4554
OY 2458 ATGCTCTCAGTATGCTGTGAGAGTCTCCAGTCCGCTGACAGACCTGATCTGAG 2517
DB 4555 ATGCTCTCAGTATGCTGTGAGAGTCTCCAGTCCGCTGACAGACCTGATCTGAG 4614
OY 2518 GACTGTGATCACAAGCCAGCGGTTTCCAGAGTCTGAGCTGAGCCCTGTCAGCAACCGG 2577
DB 4615 GACTGTGATCACAAGCCAGCGGTTTCCAGAGTCTGAGCTGAGCCCTGTCAGCAACCGG 4674
OY 2578 AGCTTGACACACCTTGTGCTTATCCAAACAAGCTGTGGGAAACGAAGTGTAAATCTACTG 2637
DB 4675 AGCTTGACACACCTTGTGCTTATCCAAACAAGCTGTGGGAAACGAAGTGTAAATCTACTG 4734
OY 2638 TGTGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATCTGATCAGTCCAG 2697
DB 4735 TGTGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATCTGATCAGTCCAG 4794
OY 2698 CTGACACAGCGCTGTGTGTTTCTTTCACCTTGTGACCTTGTGATCTGATCTGATCAG 2757
DB 4795 CTGACACAGCGCTGTGTGTTTCTTTCACCTTGTGACCTTGTGATCTGATCTGATCAG 4854
OY 2758 CACCTGACCTTATGATGAACCTCTGTGGAACAATGCGCTGGAAGCTTCTGTGAGGTC 2817
DB 4855 CACCTGACCTTATGATGAACCTCTGTGGAACAATGCGCTGGAAGCTTCTGTGAGGTC 4914
OY 2818 ATGAGAGAACAATCTTGTATCTCCAGAGACCTGAGTGTGTAAGTGTATCTTCAACGCC 2877
DB 4915 ATGAGAGAACAATCTTGTATCTCCAGAGACCTGAGTGTGTAAGTGTATCTTCAACGCC 4974
OY 2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACCTTGAAGAGCTGTGAT 2937
DB 4975 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACCTTGAAGAGCTGTGAT 5034
OY 2938 CTACAGGACAATGCGCTGTGGAACGCTGTGGTGTGCTGTGCTGTGGAAGGACTGAAGCA 2997
DB 5035 CTACAGGACAATGCGCTGTGGAACGCTGTGGTGTGCTGTGCTGTGGAAGGACTGAAGCA 5094
OY 2998 AAGAACATGTTCTGACAGAGACTGGGTTGGAAGGATGAGACTGATCTTGTGATTTGCTGT 3057
DB 5095 AAGAACATGTTCTGACAGAGACTGGGTTGGAAGGATGAGACTGATCTTGTGATTTGCTGT 5154
OY 3058 GAGGACCTCTCTTGGCCCTTCTCTGCAACCGGACTGTGACCAAGTCTTAAACCTGTGAG 3117
DB 5155 GAGGACCTCTCTTGGCCCTTCTCTGCAACCGGACTGTGACCAAGTCTTAAACCTGTGAG 5214
OY 3118 AATACTTCACTGCTCCAAAGAAATGATGAAGCTGTGTGGCTTGTGCTGTGCCAGTCT 3177
DB 5215 AATACTTCACTGCTCCAAAGAAATGATGAAGCTGTGTGGCTTGTGCTGTGCCAGTCT 5274
OY 3178 AACTTACAGATTAATTTG 3194
DB 5275 AACTTACAGATTAATTTG 5291
```

RESULT 9

AAL47131 standard; DNA; 6939 BP.

AAL47131

AAL47131;

20-AUG-2002 (first entry)

```
XX DE Pyrin domain containing protein NALPs/Py8-hs coding sequence.
XX KM Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
XX KM antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
XX KM neuroprotective; antiatheritic; antirheumatic; antiaesthetic;
XX KM nephrotoxic; osteopathic; neotropic; intracellular signal transduction;
XX KM inflammation; Alzheimer's disease; infection; psoriasis; asthma;
XX KM arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX KM osteoarthritis; glomerulonephritis; gene; ds.
XX OS Unidentified.
XX PN MO200240668-A2.
XX PD 23-MAY-2002.
XX PF 30-OCT-2001; 2001WO-EP012545.
XX PR 15-NOV-2000; 2000DE-01056687.
XX PR 30-NOV-2000; 2000DE-01059595.
XX PA (APOT-) APOTEC RES & DEV LTD.
XX PI Techopp J, Martinson F;
XX XX WPI: 2002-427093/45.
XX DR P-PSDB; AAO17859.
XX PT New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX PS Claim 5; Fig 1; 116pp; German.
XX CC The present invention relates the DNA and their encoded proteins, where
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used
XX CC to treat diseases associated with impaired intracellular signal
XX CC transduction, particularly inflammation such as psoriasis,
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX CC and Parkinson's diseases. The present sequence is a coding sequence of
XX CC the invention
XX SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX Query Match 75.8%; Score 3057.8; DB 6; Length 6939;
XX Best Local Similarity 96.2%; Pred. No. 0;
XX Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;
OY 1 ATGAGAAGAGACAATGCTTCACTTTTCCAGCTACGCGGCTGCAATGCTCTTATGAG 60
DB 3055 ATGAGAAGAGACAATGCTTCACTTTTCCAGCTACGCGGCTGCAATGCTCTTATGAG 3114
OY 61 CTAGACAAGAGAATTCAGACATTCAGGAATTTACTTAAAGAAATCTTGAGATTCG 120
DB 3115 CTAGACAAGAGAATTCAGACATTCAGGAATTTACTTAAAGAAATCTTGAGATTCG 3174
OY 121 ACCACATGCTCTATTCACAGATTGGAATGAGAAATGCAACGTCGATGCTGGCACTC 180
DB 3175 ACCACATGCTCTATTCACAGATTGGAATGAGAAATGCAACGTCGATGCTGGCACTC 3234
OY 181 CTCTTGATGAGATTTATGAGAGCTCGCTGGCTGTGCTACGTCATTTAGATTTTGA 240
DB 3235 CTCTTGATGAGATTTATGAGAGCTCGCTGGCTGTGCTACGTCATTTAGATTTTGA 3294
OY 241 AACATGAACCTGAGAACCTCTCGAGAGAGCAAGGATGACATGAAGACATTGACCA 289
DB 3295 AACATGAACCTGAGAACCTCTCGAGAGAGCAAGGATGACATGAAGACATTGACCA 3354
OY 290 ----- 289
```

Db 3355 GAAGATCTGAAGCAAGATGATGATCAAGAGCAAGAGCAAGAGAAATTCAGAGATTT 3414  
Qy 290 -----AAATTTCA 297  
Db 3415 TCACAAGCTGTGCACACAGATAGTGCACAGCTGCAGAGACAAAGAACAGAAATTTTCA 3474  
Qy 2298 CAAGCTATGGAACAAGAGGTGCACAGAGACAGAGACAGAAAGAAATTTTCA 357  
Db 3475 CAAAGCTATGGAACAAGAGGTGCACAGAGACAGAGACAGAAAGAAATTTTCA 3534  
Qy 358 GCTATGGAACAAGAGGTGCACAGAGACAGAGACAGAAAGAAATTTTCA 417  
Db 3535 GCTATGGAACAAGAGGTGCACAGAGACAGAGACAGAAAGAAATTTTCA 3594  
Qy 418 ACATGGAGATTAACAAGATGATGATGATCAAAATTTGCTGAGAGAGAGATGATGATGAT 477  
Db 3595 ACATGGAGATTAACAAGATGATGATGATCAAAATTTGCTGAGAGAGAGATGATGATGAT 3654  
Qy 478 AGTTTGAAGAACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 537  
Db 3655 AGTTTGAAGAACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3714  
Qy 538 TCAGACCGGTGGGGCTTCCGGCTGCACAGGTGTTCTGCACAGAAATTTGGG 597  
Db 3715 TCAGACCGGTGGGGCTTCCGGCTGCACAGGTGTTCTGCACAGAAATTTGGG 3774  
Qy 598 AAATGGCTCTTACCCAGAGAGATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657  
Db 3775 AAATGGCTCTTACCCAGAGAGATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3834  
Qy 658 ATGTTCTCTAGCTCTTCT 717  
Db 3835 ATGTTCTCTAGCTCTTCT 3894  
Qy 718 GTCAAGAGATTAATCTCAAGAGAGTGCAGAGCTCCAGAGCTCCGCTGACAGAGATCATG 777  
Db 3895 GTCAAGAGATTAATCTCAAGAGAGTGCAGAGCTCCAGAGCTCCGCTGACAGAGATCATG 3954  
Qy 778 TCCCAACCAAGAGAGCTGTTGTTTCAATTAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT 837  
Db 3955 TCCCAACCAAGAGAGCTGTTGTTTCAATTAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT 4014  
Qy 838 AACCAATGAACAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAG 897  
Db 4015 AACCAATGAACAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAG 4074  
Qy 898 CGCAGTCTGCTGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957  
Db 4075 CGCAGTCTGCTGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4134  
Qy 958 GTGGGACAGAGAGCTCAAGTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017  
Db 4135 GTGGGACAGAGAGCTCAAGTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4194  
Qy 1018 ATCTCCGGGGAGCAAGAGATTCATTTGCTGCTGAGCGGGAGATTTGATGATGATGATG 1077  
Db 4195 ATCTCCGGGGAGCAAGAGATTCATTTGCTGCTGAGCGGGAGATTTGATGATGATGATG 4254  
Qy 1078 ACACAAGAGTTCGCTGAGATCATGAACAACGCTGAGCTGCTGACACAGTGCAGAGTCCC 1137  
Db 4255 ACACAAGAGTTCGCTGAGATCATGAACAACGCTGAGCTGCTGACACAGTGCAGAGTCCC 4314  
Qy 1138 GCGGTGGGCTCTCTCATCTGCTGTGCGCTGCAAGCTGCAAGAGTGTGTGTGTGTGTGTGT 1197  
Db 4315 GCGGTGGGCTCTCTCATCTGCTGTGCGCTGCAAGCTGCAAGAGTGTGTGTGTGTGTGTGT 4374  
Qy 1198 GCGGTGGGCTCTCTCATCTGCTGTGCGCTGCAAGCTGCAAGAGTGTGTGTGTGTGTGTGT 1257  
Db 4375 GCGGTGGGCTCTCTCATCTGCTGTGCGCTGCAAGCTGCAAGAGTGTGTGTGTGTGTGTGT 4434  
Qy 1258 CCTGAGAGGCTGTGCTGCGGCTGTCTCAATCTGAGAGAAAGATTTCTGTAAGCGCTTC 1317  
Db 4435 CCTGAGAGGCTGTGCTGCGGCTGTCTCAATCTGAGAGAAAGATTTCTGTAAGCGCTTC 4494

Qy 1318 TCCCGTATGCTGTGAGAGAGTGTGGAATAGGAATGAGTGTGTGTGTGTGTGTGTGTGTGT 1377  
Db 4495 TCCCGTATGCTGTGAGAGAGTGTGGAATAGGAATGAGTGTGTGTGTGTGTGTGTGTGTGT 4554  
Qy 1378 ATGTTCAAGAGCTGTGGGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437  
Db 4555 ATGTTCAAGAGCTGTGGGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4614  
Qy 1438 CCAGACCCACTGTGAGAGTACTACACTTCTTCAACTTCACTGCTGCTGCTGCTGCTGCTGCT 1497  
Db 4615 CCAGACCCACTGTGAGAGTACTACACTTCTTCAACTTCACTGCTGCTGCTGCTGCTGCTGCT 4674  
Qy 1498 GCGGCTTGTACTAGTGTGTGAGAGGCTGTGGAATAGGAGCCAGCTGTGCTGCTGCTGCT 1557  
Db 4675 GCGGCTTGTACTAGTGTGTGAGAGGCTGTGGAATAGGAGCCAGCTGTGCTGCTGCTGCT 4734  
Qy 1558 GTTGAAGAACAAAGAGTTCATGAGAGCTTAAACAGCAGAGCTTCAATATCACTGCTT 1617  
Db 4735 GTTGAAGAACAAAGAGTTCATGAGAGCTTAAACAGCAGAGCTTCAATATCACTGCTT 4794  
Qy 1618 TGGATGAAGCGTTTCTGTTGTGCTGTGAGCCAGAGCTTAAGAGGCACTGAGGCTC 1677  
Db 4795 TGGATGAAGCGTTTCTGTTGTGCTGTGAGCCAGAGCTTAAGAGGCACTGAGGCTC 4854  
Qy 1678 CTGCTGGGCTGTCCCGTCTCCCTGGGGGTGAAGAGAGCTTTCACATGAGGCTCTCTG 1737  
Db 4855 CTGCTGGGCTGTCCCGTCTCCCTGGGGGTGAAGAGAGCTTTCACATGAGGCTCTCTG 4914  
Qy 1738 TTGGGTGAGAGCTTAATGCCACCACTCCAGAGAGCACTGAGAGCTTCACTGCTT 1797  
Db 4915 TTGGGTGAGAGCTTAATGCCACCACTCCAGAGAGCACTGAGAGCTTCACTGCTT 4974  
Qy 1798 TTGAGACTCAAGACAAAGATTTGTTGCTTGTGCAATTAACAGCTTCAAGAGTGTG 1857  
Db 4975 TTGAGACTCAAGACAAAGATTTGTTGCTTGTGCAATTAACAGCTTCAAGAGTGTG 5034  
Qy 1858 CTTCGATTAACCAAGACCTGAGCTTGAATAGATTTCTTCTGCTGCTGCTGCTGCTGCTG 1917  
Db 5035 CTTCGATTAACCAAGACCTGAGCTTGAATAGATTTCTTCTGCTGCTGCTGCTGCTGCTG 5094  
Qy 1918 TATTTGGGAAATTCGGGTGATGTCAAGAGATTTCTCCAGAGATGAGTCCGCTGAG 1977  
Db 5095 TATTTGGGAAATTCGGGTGATGTCAAGAGATTTCTCCAGAGATGAGTCCGCTGAG 5154  
Qy 1978 GCATGCTCTGTGTCTCTCTATGATGCGGATTAAGACCTCATTTGAGAGACAGTGGAA 2037  
Db 5155 GCATGCTCTGTGTCTCTCTATGATGCGGATTAAGACCTCATTTGAGAGACAGTGGAA 5214  
Qy 2038 GATTTCTGCTCAATGCTTTGGCACTCACTGCGGAGCTGAGCTTGGGAGCAGC 2097  
Db 5215 GATTTCTGCTCAATGCTTTGGCACTCACTGCGGAGCTGAGCTTGGGAGCAGC 5274  
Qy 2098 ATCTGAGAGAGCGGGACATGAAGACCTGTGTGCAAGCTGAGAGCATCCACCTGCAAG 2157  
Db 5275 ATCTGAGAGAGCGGGACATGAAGACCTGTGTGCAAGCTGAGAGCATCCACCTGCAAG 5334  
Qy 2158 ATACAGACCTGATGTTTGAAGATGCAAGATTAACCTCTGCTGTGCAAGCTTGTGAG 2217  
Db 5335 ATACAGACCTGATGTTTGAAGATGCAAGATTAACCTCTGCTGTGCAAGCTTGTGAG 5394  
Qy 2218 ATGCTATGCGCAACCTTAACCTTAAGATCCTTCAACTTTGGAGGACCCACCTGAAGAA 2277  
Db 5395 ATGCTATGCGCAACCTTAACCTTAAGATCCTTCAACTTTGGAGGACCCACCTGAAGAA 5454  
Qy 2278 GAGAGTGAAGAGAGGCTGTAAGAGCTTAAACACCAAAATTTGTGTGAGCTTGTG 2337  
Db 5455 GAGAGTGAAGAGAGGCTGTAAGAGCTTAAACACCAAAATTTGTGTGAGCTTGTG 5514  
Qy 2338 AGGCTGATGCTGTGATGATGACCATGCTGTTACCTGAAGATCTCCAAATCTTAAG 2397  
Db 5515 AGGCTGATGCTGTGATGATGACCATGCTGTTACCTGAAGATCTCCAAATCTTAAG 5574



```

Oy 2398 ACCTCCCGCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAAGTGACAGACCAGGAGTA 2457
    |||||
Db 5575 ACCCTCCCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAAGTGACAGACCAGGAGTA 5634
Oy 2458 ATGCTCTCAGTGATGCTTGAAGTCTTCCAGTGGCCCTGACAGAGCTGATCTGGAG 2517
    |||||
Db 5635 ATGCTCTCAGTGATGCTTGAAGTCTTCCAGTGGCCCTGACAGAGCTGATCTGGAG 5694
Oy 2518 GACTGGGATCAGACGCGGCTTGGCCAGAGTCTGGCCTGAGCCCTGGTGCAGAACCGG 2577
    |||||
Db 5695 GACTGGGATCAGACGCGGCTTGGCCAGAGTCTGGCCTGAGCCCTGGTGCAGAACCGG 5754
Oy 2578 AGCTTGACACACCTTGCTGCTTACCAACACAGCCTGGGAAAGAAAGTGAATCTACTG 2637
    |||||
Db 5755 AGCTTGACACACCTTGCTGCTTACCAACACAGCCTGGGAAAGAAAGTGAATCTACTG 5814
Oy 2638 TGTGCATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAGTCCAC 2697
    |||||
Db 5815 TGTGCATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAGTCCAC 5874
Oy 2698 CTGACACGCGCTGCTGTGTTTCTTGCACTTGGCTTATGGGTAACTCAATGGCTGACG 2757
    |||||
Db 5875 CTGACACGCGCTGCTGTGTTTCTTGCACTTGGCTTATGGGTAACTCAATGGCTGACG 5934
Oy 2758 CACCTGAGCCTTGAAGTGAACCCCTGTGAAGACAAATGGCGTGAAGCTTCTGTGCGAGGTC 2817
    |||||
Db 5935 CACCTGAGCCTTGAAGTGAACCCCTGTGTGAAGACAAATGGCGTGAAGCTTCTGTGCGAGGTC 5994
Oy 2818 ATGAGAAACCATCTTGTTCATCTCCAGAGCCTGGAGTTGGTAAAGTGTCTCATCTCCGCC 2877
    |||||
Db 5995 ATGAGAAACCATCTTGTTCATCTCCAGAGCCTGGAGTTGGTAAAGTGTCTCATCTCCGCC 6054
Oy 2878 GCGTCTGAGAGTCTGTCTGCTGTGTGATCTGAGAGAGACACCTGAAGAGCTTGAT 2937
    |||||
Db 6055 GCGTCTGAGAGTCTGTCTGCTGTGTGATCTGAGAGAGACACCTGAAGAGCTTGAT 6114
Oy 2938 CTGACGGAACATGCTGCTGGTGAACCGTGGGCTTGTGCTGCTGTGCGAGGACCTGAAGCA 2997
    |||||
Db 6115 CTGACGGAACATGCTGCTGGTGAACCGTGGGCTTGTGCTGCTGTGCGAGGACCTGAAGCA 6174
Oy 2998 AAGAACAGTGTCTGACGAGACTCGGCTTGAAGGACATGTGACTGATCTTGATTTGCTGT 3057
    |||||
Db 6175 AAGAACAGTGTCTGACGAGACTCGGCTTGAAGGACATGTGACTGATCTTGATTTGCTGT 6234
Oy 3058 GAGGCACTCTCTTGGCCCTTCCGCAACCGGACATGACCAAGCTAAACCTGGTGGAG 3117
    |||||
Db 6235 GAGGCACTCTCTTGGCCCTTCCGCAACCGGACATGACCAAGCTAAACCTGGTGGAG 6294
Oy 3118 AATTAATTCACTTCCAAAGAAATGATGAAGCTGTGCTGGCTTGTGCTGCTCCACGCT 3177
    |||||
Db 6295 AATTAATTCACTTCCAAAGAAATGATGAAGCTGTGCTGGCTTGTGCTGCTCCACGCT 6354
Oy 3178 AACTTACGATTAATGG 3194
    |||||
Db 6355 AACTTACGATTAATGG 6371
    |||||

```

## RESULT 10

AA147140  
ID AA147140 standard; DNA, 6939 BP.

AA147140;

20-AUG-2002 (first entry)

Pyrin domain containing protein NALP3/Py17 coding sequence.

XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;  
XX antiarteriosclerotic; antipsoriatic; antibacterial; varicoid;  
XX neuroprotective; antiarthritic; antirheumatic; antiashtmatic;  
XX nephrotropic; osteopathic; nootropic; intracellular signal transduction;  
XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;

```

KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS unidentified.
XX
XX WO200240668-A2.
XX
XX 23-MAY-2002.
XX
XX 30-OCT-2001; 2001WO-BP012545.
XX
XX 15-NOV-2000; 2000DE-01056687.
XX
XX 30-NOV-2000; 2000DE-01059595.
XX
XX (APOT-) APOTEC RES & DEV LTD.
XX
XX Tschoep J, Martinson F;
XX
XX WPI; 2002-427093/45.
XX
XX P-PSDB; MA017868.
XX
XX
XX New DNA encoding protein with pyrin domain, useful for treating diseases
XX involving impaired signal transduction, particularly inflammation, also
XX proteins and antibodies.
XX
XX Claim 5; Fig 1; 116pp; German.
XX
XX The present invention relates the DNA and their encoded proteins, where
XX the proteins contain at least one PYD (pyrin) domain. These can be used
XX to treat diseases associated with impaired intracellular signal
XX transduction, particularly inflammation such as psoriasis,
XX arteriosclerosis, bacterial or viral infections (particularly meningitis
XX and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX the invention
XX
XX Sequence 6939 BP; 1791 A; 1719 G; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX Query Match 75.8%; Score 3057.8; DB 6; Length 6939;
XX Best Local Similarity 96.2%; Pred. No. 0;
XX Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;
Oy 1 ATGAGAGAGACAAATGCTCACCCTTTCAGCTACGCGGCTGCATGCTGTCTATGAG 60
    |||||
Db 3055 ATGAGAGAGACAAATGCTCACCCTTTCAGCTACGCGGCTGCATGCTGTCTATGAG 3114
Oy 61 CTAGACAAAGAAATTTCAAGACTTCAAGAAATTAATAAGAAATTTCAAGAAATCG 120
    |||||
Db 3115 CTAGACAAAGAAATTTCAAGACTTCAAGAAATTAATAAGAAATTTCAAGAAATCG 3174
Oy 121 ACCACATGCTCTATTCCACAGTTTGAATGCAAGATGCAACGTAAGTGTGGCACTC 180
    |||||
Db 3175 ACCACATGCTCTATTCCACAGTTTGAATGCAAGATGCAACGTAAGTGTGGCACTC 3234
Oy 181 CTCTTGATGATGATTTATGAGATGCTGCGCTGGCTGACGTCATTAATGATCTTTGAA 240
    |||||
Db 3235 CTCTTGATGATGATTTATGAGATGCTGCGCTGGCTGACGTCATTAATGATCTTTGAA 3294
Oy 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGCACGGATGACATGAATAA----- 289
    |||||
Db 3295 AACATGAACCTGCGCAACCTCTCGGAGAAAGCACGGATGACATGAATAAAGCATTCACA 3354
    |||||
Oy 290 ----- 289
    |||||
Db 3355 GAAGATCCTGAAGCAAGATGACTGACCAAGAACCAAGAAAGTCCAGGAATT 3414
    |||||
Oy 290 -----AATTTCA 297
    |||||
Db 3415 TCACAAGCTGCAACAAGATAGTGCACAGCTGCAGAGACAAAGAAAGAAATTTCA 3474
    |||||
Oy 298 CAAGCTATGAAACAAGAGGTGCCACAGCAGCAGAGACAGAGAAACAAGAAATTTCA 357
    |||||
Db 3475 CAAGCTATGAAACAAGAGGTGCCACAGCAGCAGAGACAGAGAAACAAGAAATTTCA 3534
    |||||

```



358 GGTATGAAACAAGAGTGGCAAGAGAGAGAGAAACAAGATGAGAGTAC 417  
3535 GCTATGAAACAAGAGTGGCAAGAGAGAGAGAAACAAGATGAGAGTAC 3594  
418 ACATGGAGCTACAGAGTACGATGACAAATTCGCTGAGAGAGAGATGACCTGCT 477  
3595 ACATGGAGCTACAGAGTACGATGACAAATTCGCTGAGAGAGAGATGACCTGCT 3654  
478 AGTTTGAACAACCTGCTGATGACCTGAGAAATGCAACCTTGGCTGCTTTTAT 537  
3655 AGTTTGAACAACCTGCTGATGACCTGAGAAATGCAACCTTGGCTGCTTTTAT 3714  
538 TCAGACCGGTGGGGCTTCCGGCTGGCAAGGTGTTCTGCAAGAAAGTCAAGAAATGGG 597  
3715 TCAGACCGGTGGGGCTTCCGGCTGGCAAGGTGTTCTGCAAGAAAGTCAAGAAATGGG 3774  
598 AATCGGCTCTAGCCAGAGAGATCGTGTGCTGGCGCAAGAGTGAAGCTTACAGAGA 657  
3775 AATCGGCTCTAGCCAGAGAGATCGTGTGCTGGCGCAAGAGTGAAGCTTACAGAGA 3834  
658 ATGTTCTCTACGCTTCTTCTTCTCCCGTTAAGAGATGACAGCGAAGAGAGACAT 717  
3835 ATGTTCTCTACGCTTCTTCTTCTCCCGTTAAGAGATGACAGCGAAGAGAGACAT 3894  
718 GTACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGAGCTCCGGTGAAGAGATCATG 777  
3895 GTACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGAGCTCCGGTGAAGAGATCATG 3954  
778 TCCCGACAGAAAGGCTGTGTTTCATCATTTGACGGTTTGCATGACCTGGGCTCTGCTC 837  
3955 TCCCGACAGAAAGGCTGTGTTTCATCATTTGACGGTTTGCATGACCTGGGCTCTGCTC 4014  
838 AACATGACACAAGCTCTGCAAAAGCTGGCTGAGAGACAGCTCCGTTCAACCTCAT 897  
4015 AACATGACACAAGCTCTGCAAAAGCTGGCTGAGAGACAGCTCCGTTCAACCTCAT 4074  
898 CGCAGTCTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACCTGAGAGAC 957  
4075 CGCAGTCTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACCTGAGAGAC 4134  
958 GTGGGACAGAGAGTCAAGTCAAGAGTGTGCTCCCGTTACCTGTAGTTAGAGA 1017  
4135 GTGGGACAGAGAGTCAAGTCAAGAGTGTGCTCCCGTTACCTGTAGTTAGAGA 4194  
1018 ATTCCTGGGAAACAAGATTCACCTTGTCTTGAAGCGGAGATTGGTAGATCAGAG 1077  
4195 ATTCCTGGGAAACAAGATTCACCTTGTCTTGAAGCGGAGATTGGTAGATCAGAG 4254  
1078 ACACAGAGGTTGCTGCTGATCATGAAACAAGCTGAGCTGTCGACAGAGTGGCC 1137  
4255 ACACAGAGGTTGCTGCTGATCATGAAACAAGCTGAGCTGTCGACAGAGTGGCC 4314  
1138 GCGGTGGGCTCTCATCTGCGGTGGCCCTGAGCTGAGAGAGTGGTGGGGAGAGCTC 1197  
4315 GCGGTGGGCTCTCATCTGCGGTGGCCCTGAGCTGAGAGAGTGGTGGGGAGAGCTC 4374  
1198 GCGCCCTTCAACAACGCTCAAGAGCTGACAGCGCTTTTGTGTTTCAATGACTCAC 1257  
4375 GCGCCCTTCAACAACGCTCAAGAGCTGACAGCGCTTTTGTGTTTCAATGACTCAC 4434  
1258 CTTGAGAGGCTGCTGCGGCTGCTGCTCAATCTGAGAGAGAGATTGCTGAGAGCTTC 1317  
4435 CTTGAGAGGCTGCTGCGGCTGCTGCTCAATCTGAGAGAGAGATTGCTGAGAGCTTC 4494  
1318 TGCCTGATGCTGTGAGAGGAGTGAATAGAAAGTCAAGTTTGAAGTGAAGAGCTC 1377  
4495 TGCCTGATGCTGTGAGAGGAGTGAATAGAAAGTCAAGTTTGAAGTGAAGAGCTC 4554  
1378 ATGTTTCAAGAGCTCGGGAGTCTGAGCTCCGTGCTGTGTTTCAATGAAATCTTCTC 1437  
4555 ATGTTTCAAGAGCTCGGGAGTCTGAGCTCCGTGCTGTGTTTCAATGAAATCTTCTC 4614

1438 CCAGACAGCTGTGAGAGTACTACCTTCTTCAACCTGACTCAGAGACTTCTGT 1497  
4615 CCAGACAGCTGTGAGAGTACTACCTTCTTCAACCTGACTCAGAGACTTCTGT 4674  
1498 GCGGCTTGTACTAGTGTGAGAGGCTGGAATGAGAGAGCTCTGTGCTTGTAC 1557  
4675 GCGGCTTGTACTAGTGTGAGAGGCTGGAATGAGAGAGCTCTGTGCTTGTAC 4734  
1558 GTTGAAGACAAAGAGTCCATGAGAGTAAACAGAGAGCTTCAATCACTGCTT 1617  
4735 GTTGAAGACAAAGAGTCCATGAGAGTAAACAGAGAGCTTCAATCACTGCTT 4794  
1618 TGATGAAGAGCTTTTGTGAGAGCTGAGAGAGAGTAAAGAGAGCTTCAAGAGCTC 1677  
4795 TGATGAAGAGCTTTTGTGAGAGCTGAGAGAGAGTAAAGAGAGCTTCAAGAGCTC 4854  
1678 CTGCTGGGCTGCTCCGTTCCCTGGGGGTGAAGAGAGCTTTCAGACTGCTCTG 1737  
4855 CTGCTGGGCTGCTCCGTTCCCTGGGGGTGAAGAGAGCTTTCAGACTGCTCTG 4914  
1738 TTGGGTACAGAGCTTAATGCAACACCCAGAGAGACCTTGAAGCTTCACTGCTT 1797  
4915 TTGGGTACAGAGCTTAATGCAACACCCAGAGAGACCTTGAAGCTTCACTGCTT 4974  
1798 TTGGGTACAGAGCTTAATGCAACACCCAGAGAGACCTTGAAGCTTCACTGCTT 1857  
4975 TTGGGTACAGAGCTTAATGCAACACCCAGAGAGACCTTGAAGCTTCACTGCTT 5034  
1858 CTTCCGATTAACAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATG 1917  
5035 CTTCCGATTAACAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATG 5094  
1918 TATTTGCGAAATTCGGGTGATGATGATGATGATGATGATGATGATGATGATG 1977  
5095 TATTTGCGAAATTCGGGTGATGATGATGATGATGATGATGATGATGATGATG 5154  
1978 GCATGCTCTGAGTCCCTCTATGATGATGATGATGATGATGATGATGATGATG 2037  
5155 GCATGCTCTGAGTCCCTCTATGATGATGATGATGATGATGATGATGATGATG 5214  
2038 GATTTCTGCTCATGCTTGGAGACCAACCACTGCGGAGCTGAGACTGGGAGCAGC 2097  
5215 GATTTCTGCTCATGCTTGGAGACCAACCACTGCGGAGCTGAGACTGGGAGCAGC 5274  
2098 ATCTGACAGAGCGGAGCATGAAGACCTGTGTGCAAGCTGAGAGATCCACTGCAAG 2157  
5275 ATCTGACAGAGCGGAGCATGAAGACCTGTGTGCAAGCTGAGAGATCCACTGCAAG 5334  
2158 ATACAGAGCTGATGTTTGAAGATGACAGATTAACCTGCTGCTGAGAGACTTGGAGA 2217  
5335 ATACAGAGCTGATGTTTGAAGATGACAGATTAACCTGCTGCTGAGAGACTTGGAGA 5394  
2218 ATTCATGAGGCAACCGTAACTTAAGATCCCTCAACTTGGAGAGCAACCTGAAGAA 2277  
5395 ATTCATGAGGCAACCGTAACTTAAGATCCCTCAACTTGGAGAGCAACCTGAAGAA 5454  
2278 GAGAGTGAAGAGTGGCTGTGAAGCTTAAACCAACCAAAATGTTTGGAGTCTTTG 2337  
5455 GAGAGTGAAGAGTGGCTGTGAAGCTTAAACCAACCAAAATGTTTGGAGTCTTTG 5514  
2338 AGGCTGATGCTGTGATGATGACCAATGCTGTTAAGTGAAGTCTCCAAATCTTAAG 2397  
5515 AGGCTGATGCTGTGATGATGACCAATGCTGTTAAGTGAAGTCTCCAAATCTTAAG 5574  
2398 ACTTCCCGACCTGAATCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2457  
5575 ACTTCCCGACCTGAATCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 5634  
2458 ATGCTCTCAATGATGCTTGAAGAGTCTCCAGTGGCGCTTGAAGAGCTGATCTGAG 2517  
5635 ATGCTCTCAATGATGCTTGAAGAGTCTCCAGTGGCGCTTGAAGAGCTGATCTGAG 5694  
2518 GACTGTGATCAACAGCAGCGGTTGCAAGAGTCTGAGCTTGAAGAGCTTGAAGAG 2577

Db 5695 GACTGTGGCATCAGACCAAGGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACGG 5754  
Qy 2578 AGCTTGACACACCTGTGCTATCCAAACAAGCCCTGGGAAAGAGGTGAATCTACTG 2637  
Db 5755 AGCTTGACACACCTGTGCTATCCAAACAAGCCCTGGGAAAGAGGTGAATCTACTG 5814  
Qy 2638 TGTGCATCATAGAGCTTCCCACTGATGTCTGACAGAGGCTGATCTGAATCAGTCCAC 2697  
Db 5815 TGTGCATCATAGAGCTTCCCACTGATGTCTGACAGAGGCTGATCTGAATCAGTCCAC 5874  
Qy 2698 CTGACACGGCTGCTGTGTCTTCTTGCACTTGGCTTATGGTAACTCATGCTGACG 2757  
Db 5875 CTGACACGGCTGCTGTGTCTTCTTGCACTTGGCTTATGGTAACTCATGCTGACG 5934  
Qy 2758 CACCGAGCCTTAGATGAACCTGTGAAAGCAATGACGTGAACCTTCTGTGCGAGGTC 2817  
Db 5935 CACCGAGCCTTAGATGAACCTGTGAAAGCAATGACGTGAACCTTCTGTGCGAGGTC 5994  
Qy 2818 ATGAGAGAACCATCTTGTATCTCCAGGACCTGAGTTGTAAAGTGTATCTACCGCC 2877  
Db 5995 ATGAGAGAACCATCTTGTATCTCCAGGACCTGAGTTGTAAAGTGTATCTACCGCC 6054  
Qy 2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTCCAGGACCAACACCTGAAGGCTTGAT 2937  
Db 6055 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTCCAGGACCAACACCTGAAGGCTTGAT 6114  
Qy 2938 CTGACGGAACAAGCCCTGGTGAACGCTGGGGTGTCTGGCTGTGAGAGGACTGAAGAA 2997  
Db 6115 CTGACGGAACAAGCCCTGGTGAACGCTGGGGTGTCTGGCTGTGAGAGGACTGAAGAA 6174  
Qy 2998 AAGAACAGTGTCTGACGAGACTCGGGTGAAGGATGAGACTGACTTCTGATGTCTGT 3057  
Db 6175 AAGAACAGTGTCTGACGAGACTCGGGTGAAGGATGAGACTGACTTCTGATGTCTGT 6234  
Qy 3058 GAGGCACTCTCTTGGCCCTTCTGCAACGGGCACTGACCAAGTCTAAACCTGTGAG 3117  
Db 6235 GAGGCACTCTCTTGGCCCTTCTGCAACGGGCACTGACCAAGTCTAAACCTGTGAG 6294  
Qy 3118 AATACTCAATGCCAAAGGAATGATGAAGCTGTGGCCTTGGCCCTGCCACGCT 3177  
Db 6295 AATACTCAATGCCAAAGGAATGATGAAGCTGTGTGGCCTTGGCCCTGCCACGCT 6354  
Qy 3178 AACTTACAGATTAATGG 3194  
Db 6355 AACTTACAGATTAATGG 6371

RESULT 11  
ABX97181  
ID ABX97181 standard; cDNA; 3226 BP.  
AC ABX97181;  
XX  
XX 20-MAY-2003 (first entry)  
DT XX  
DE Human NOV125b cDNA.  
XX  
XX NOX; cytoelastic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;  
KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
KM human; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200272757-A2.  
PN  
XX 19-SEP-2002.  
PD  
XX  
XX 08-MAR-2002; 2002WO-US006908.  
PF  
XX  
XX 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR  
XX

PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0277338P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 07-MAR-2002; 2002US-00092900.  
PR  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;  
PI Zethusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Paturajan M, Gangoli E, Vernet CM, Guo X, Tchernev V;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;  
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsbrook JP;  
PI Lepley DM, Rieger DK;  
DR WPI; 2002-72332/78.  
DR P-PSDB; ABU65214.  
DR  
XX

PT NOVX polypeptides and polynucleotides, useful for preventing or treating  
PT a disorder associated with aberrant NOVX expression or activity e.g.,  
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
PT asthma.

XX Claim 13; Page 467; 1103bp; English.

CC This invention describes novel human NOVX polypeptides which have  
CC cytoskeletal, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive  
CC activity. Pharmaceutical compositions comprising the NOVX proteins or  
CC nucleic acid molecules or NOVX antibodies are useful for preventing or  
CC treating a disorder associated with aberrant NOVX expression or activity  
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
CC asthma. The products of the invention can be used for gene therapy or in  
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by  
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-  
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in  
CC ABU65041-ABU65218

XX Sequence 3226 BP; 738 A; 856 C; 914 G; 718 T; 0 U; 0 Other;

Query Match 71.9%; Score 2900.6; DB 6; Length 3226;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2903; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	406	CATGGAGGAGACATGGGACCTACAAAGATCAAGTACGATGACCAATTGCGTAAAGAGAG	465
DB	249	CATGGAGGAGACATGGGACCTACAAAGATCAAGTACGATGACCAAAATGCTGTAAGAGAGAG	308
QY	466	GATGATACGTCGTAGTTTGAACCACTGCTGCTGACTGGCCGGAATGCAACGTTGGCT	525
DB	309	GATGATACGTCGTAGTTTGAACCACTGCTGCTGACTGGCCGGAATGCAACGTTGGCT	368
QY	526	GCTGCTTTGATTCAGACCGGTGGGCTTCGCGCTCCGACGCTGGTTCGACGGAAG	585
DB	369	GCTGCTTTGATTCAGACCGGTGGGCTTCGCGCTCCGACGCTGGTTCGACGGAAG	428
QY	586	TCAGGAATTTGGGAAATCGGCTTACGCAAGAGATCGTGTGCTGGGCGCAAGTGA	645
DB	429	TCAGGAATTTGGGAAATCGGCTTACGCAAGAGATCGTGTGCTGGGCGCAAGTGA	488
QY	646	CTCTACGAGGAAATGTTCTCTACGCTCTTCTCCCGTTAGAGAGATGAGGGAAG	705
DB	489	CTCTACGAGGAAATGTTCTCTACGCTCTTCTCCCGTTAGAGAGATGAGGGAAG	548
QY	706	AAGAGAGACGATGTCACAGAGTTTCTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTG	765
DB	549	AAGAGAGACGATGTCACAGAGTTTCTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTG	608
QY	766	ACGGAGATCATGTCCCGACACAGAAAGGCTGTTTCTATCATTTAGAGGTTTGAATGACTG	825
DB	609	ACGGAGATCATGTCCCGACACAGAAAGGCTGTTTCTATCATTTAGAGGTTTGAATGACTG	668
QY	826	GCTCTGTCCTCAACATGACCAAGAGCTCTGCAAAAGCTGGGCTGAGAGAGGCTCCG	885
DB	669	GCTCTGTCCTCAACATGACCAAGAGCTCTGCAAAAGCTGGGCTGAGAGAGGCTCCG	728
QY	886	TTTCACTCTCATACGAGTCTGCTGAGAGAGCTCTGCTCCCTGAGTCTCTTCTGATCTGTC	945
DB	729	TTTCACTCTCATACGAGTCTGCTGAGAGAGCTCTGCTCCCTGAGTCTCTTCTGATCTGTC	788
QY	946	ACCGTCAGAGAGTGGGACACAGAAAGCTCAAGTCAAGAGTGTGTCTTCCCGTTACTGTC	1005
DB	789	ACCGTCAGAGAGTGGGACACAGAAAGCTCAAGTCAAGAGTGTGTGTCTTCCCGTTACTGTC	848
QY	1006	TTAGTTAGAGAAATCTCCGGGGAGCAAAAGATCCACTTGCTCTTGAAGCGGGATTGGT	1065
DB	849	TTAGTTAGAGAAATCTCCGGGGAGCAAAAGATCCACTTGCTCTTGAAGCGGGATTGGT	908
QY	1066	GAGCATCAGAGACCAAGGGTGGTGGTGGATCATGAACCAACCGTAGGCTGCTCAGCAG	1125
DB	909	GAGCATCAGAGACCAAGGGTGGTGGTGGATCATGAACCAACCGTAGGCTGCTCAGCAG	968

QY	1126	TGCCAGTGCCTCCGCGGCTCTCTCATCTGCGTGGCCCTGACAGTCCAGAGCTGGTG	1185
DB	969	TGCCAGTGCCTCCGCGGCTCTCTCATCTGCGTGGCCCTGACAGTCCAGAGCTGGTG	1028
QY	1186	GGGAGAGCGTGGCCCTTCAACCAACGCTCAAGGCTTCAGCGCGCTTTTGTGTTT	1245
DB	1029	GGGAGAGCGTGGCCCTTCAACCAACGCTCAAGGCTTCAGCGCGCTTTTGTGTTT	1088
QY	1246	CATGATCTACCCCTTCAGAGGCTGTCCGGGCTGTCTCATCTTGGAGAGAAAGATGTGTC	1305
DB	1089	CATGATCTACCCCTTCAGAGGCTGTCCGGGCTGTCTCATCTTGGAGAGAAAGATGTGTC	1148
QY	1306	CTGAAGCGCTTCGCGGATGATGGCTGTGAGAGGAGTGGAAATGGAAGTCAAGTGTTCAT	1365
DB	1149	CTGAAGCGCTTCGCGGATGATGGCTGTGAGAGGAGTGGAAATGGAAGTCAAGTGTTCAT	1208
QY	1366	GATGACGACTCATAGTTTCAAGAGACTCGGAGAGTCTGAGTCCGTCTGTGTTACATG	1425
DB	1209	GATGACGACTCATAGTTTCAAGAGACTCGGAGAGTCTGAGTCCGTCTGTGTTACATG	1268
QY	1426	AACATCTTCTTCCAGACAGCCACTGTGAGAGATCTACACTTCTTCACTCAGTCTC	1485
DB	1269	AACATCTTCTTCCAGACAGCCACTGTGAGAGATCTACACTTCTTCACTCAGTCTC	1328
QY	1486	CAGGACTTCTGAGCGCTTGTACTAGTGTAGAGGCGCTGGAATGAGCCAGCTCTC	1545
DB	1329	CAGGACTTCTGAGCGCTTGTACTAGTGTAGAGGCGCTGGAATGAGCCAGCTCTC	1388
QY	1546	TGCCCTCTGTACGTTGAGAGACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCAT	1605
DB	1389	TGCCCTCTGTACGTTGAGAGACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCAT	1448
QY	1606	ATCCACTGCTTTGGATGATTAACGCTTCTGTTTGGCTCTGTAGAGAGAGATGAGAGG	1665
DB	1449	ATCCACTGCTTTGGATGATTAACGCTTCTGTTTGGCTCTGTAGAGAGAGATGAGAGG	1508
QY	1666	COACTGAGAGTCTGCTGGGCTGTCCGCTCCCGGGGGGAGAGCAAGCTTCTGAC	1725
DB	1509	COACTGAGAGTCTGCTGGGCTGTCCGCTCCCGGGGGGAGAGCAAGCTTCTGAC	1568
QY	1726	TGGGCTCTCTGTTGGGTCAGACAGCTTAATGCAACACCCAGAGAGACACCTGAGACGCC	1785
DB	1569	TGGGCTCTCTGTTGGGTCAGACAGCTTAATGCAACACCCAGAGAGACACCTGAGACGCC	1628
QY	1786	TTTCACTGTCTTTTGAAGACTCAAGACAAAGATTGTTGGCTTGGCAATTAAACAGCTTC	1845
DB	1629	TTTCACTGTCTTTTGAAGACTCAAGACAAAGATTGTTGGCTTGGCAATTAAACAGCTTC	1688
QY	1846	CAAGAGTGTGGCTTCCGATTAACCAAGAACTTGAACCTTGAATGATCTTCTTCTGCTC	1905
DB	1689	CAAGAGTGTGGCTTCCGATTAACCAAGAACTTGAACCTTGAATGATCTTCTTCTGCTC	1748
QY	1906	CAGACGTGCTCGTATTGTCGGAATAATTCGGGTGATGCAAAAGGATCTTCCCAAGAGAT	1965
DB	1749	CAGACGTGCTCGTATTGTCGGAATAATTCGGGTGATGCAAAAGGATCTTCCCAAGAGAT	1808
QY	1966	GAGTCCGCTGAGGAGATCTGTGCTCTCTATGATGCGGATTAAGACCTCATTTAG	2025
DB	1809	GAGTCCGCTGAGGAGATCTGTGCTCTCTATGATGCGGATTAAGACCTCATTTAG	1868
QY	2026	GAGGATGGGAAATTTCTGCTCATAGCTTGGACCCACCACTTCGCGGACCTGGAC	2085
DB	1869	GAGGATGGGAAATTTCTGCTCATAGCTTGGACCCACCACTTCGCGGACCTGGAC	1928
QY	2086	CTGGGACAGACATCTTGAACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGCAT	2145
DB	1929	CTGGGACAGACATCTTGAACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGCAT	1988
QY	2146	CCCACTTGAAGATACAGACCTTGATGTTTGAAGATCAAGATTAACCTTGGTGTGAG	2205
DB	1989	CCCACTTGAAGATACAGACCTTGATGTTTGAAGATCAAGATTAACCTTGGTGTGAG	2048
QY	2206	CACCTTGGAGAAATGTCATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGGACCC	2265

Db 2049 CACCTCTGGAGATCGTCATGGCCAAACGTAACCTAAGATCCCTCACTCTGGAGGACACC 2108  
Qy 2266 CACCTGAAGAAAGAGATGTAAGATGGCGCTGTGAAGCCTTAAACACCCAAAAGTTTG 2325  
Db 2109 CACCTGAAGAAAGAGATGTAAGATGGCGCTGTGAAGCCTTAAACACCCAAAAGTTTG 2168  
Qy 2326 TTGAGATCTTTGAGGCTGGAATTTGCTGTGGAATGACCCATGCGTGTACTGTGAAGATCTCC 2385  
Db 2169 TTGAGATCTTTGAGGCTGGAATTTGCTGTGGAATGACCCATGCGTGTACTGTGAAGATCTCC 2228  
Qy 2386 CAATCTTTAGACCTCCCGACCTGAAATCTGTAGCCTTGGCAGAAACAAGTGAACA 2445  
Db 2229 CAATCTTTAGACCTCCCGACCTGAAATCTGTAGCCTTGGCAGAAACAAGTGAACA 2288  
Qy 2446 GACCGAGGAGTAATGCTCTGATGATGCTTTGAGATCTCCCAATGCGCCCTGAGAG 2505  
Db 2289 GACCGAGGAGTAATGCTCTGATGATGCTTTGAGATGCTCCCAATGCGCCCTGAGAG 2348  
Qy 2506 CTGATCTGAGAGATCTGTGATCAACAGCCAGGTTCCAGAGTCTGGCCTCAGCCCTC 2565  
Db 2349 CTGATCTGAGAGATCTGTGATCAACAGCCAGGTTCCAGAGTCTGGCCTCAGCCCTC 2408  
Qy 2566 GTGACCAACCGAGCTTGAACAACCTGTGCTATCAACAACAGCTGAGGAAACGAAAGT 2625  
Db 2409 GTGACCAACCGAGCTTGAACAACCTGTGCTATCAACAACAGCTGAGGAAACGAAAGT 2468  
Qy 2626 GTAAATCTACTGTGTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2685  
Db 2469 GTAAATCTACTGTGTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2528  
Qy 2686 AATCAGTCCCACTGTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2745  
Db 2529 AATCAGTCCCACTGTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2588  
Qy 2746 TCATGCTGAGGATCTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2805  
Db 2589 TCATGCTGAGGATCTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2648  
Qy 2806 CTGTGCGAGGATCTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2865  
Db 2649 CTGTGCGAGGATCTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2708  
Qy 2866 CATCTCACCGCGCTGTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2925  
Db 2709 CATCTCACCGCGCTGTGATGATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTG 2768  
Qy 2926 AAGAGCTGTGATCTCAACGACATGCTGAGGCTTCCCACTGTAGTGTGCAAGAGCTG 2985  
Db 2769 AAGAGCTGTGATCTCAACGACATGCTGAGGCTTCCCACTGTAGTGTGCAAGAGCTG 2828  
Qy 2986 GGACTGAGGAAAGAAAGATGTTCTGAAGACTCGGAGTTGAAGGCAATGTGAGACTG 3045  
Db 2829 GGACTGAGGAAAGAAAGATGTTCTGAAGACTCGGAGTTGAAGGCAATGTGAGACTG 2888  
Qy 3046 TCTGATCTGTGAGGCACTCTCTTGAGCCTTTCTGCAACCGGATCTGACCACTG 3105  
Db 2889 TCTGATCTGTGAGGCACTCTCTTGAGCCTTTCTGCAACCGGATCTGACCACTG 2948  
Qy 3106 AACCTGTGAGGATCTTCAAGTCCCAAGAGATGATGAAAGCTGTGTTGGCCTTTGCC 3165  
Db 2949 AACCTGTGAGGATCTTCAAGTCCCAAGAGATGATGAAAGCTGTGTTGGCCTTTGCC 3008  
Qy 3166 TGTCCAGCTCTAATTAATTAATTAATTTGGGCTGTGGAATGAGCAATACCTGTGCAATA 3225  
Db 3009 TGTCCAGCTCTAATTAATTAATTAATTTGGGCTGTGGAATGAGCAATACCTGTGCAATA 3068  
Qy 3226 AGAAGAGCTGTGAGGAGATGAGTACTCAAGCCCGAGTGTGTAATGAGGATGTTGG 3285  
Db 3069 AGAAGAGCTGTGAGGAGATGAGTACTCAAGCCCGAGTGTGTAATGAGGATGTTGG 3128  
Qy 3286 CATCTTTTGAATGAAGATGACCGACAC 3312

Db 3129 CATCTTTTGAATGAAGATGACCGGTAC 3155  
RESULT 12  
ID ADN62078 standard; cDNA; 2820 BP.  
XX ADN62078;  
AC  
XX 01-JUL-2004 (first entry)  
DT  
XX  
DE Human cDNA encoding NOV125a.  
XX  
XX Human; ss; gene; NOXV; diabetes; obesity; infectious disease; anorexia;  
KW cancer-associated cachexia; cancer; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; chronic disease.  
XX  
OS Homo sapiens.  
XX  
XX US2004043382-A1.  
PN  
XX 04-MAR-2004.  
PD  
XX  
XX 07-MAR-2002; 2002US-00092900.  
PF  
XX  
XX 08-MAR-2001; 2001US-0274191P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277338P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278884P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279959P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281444P.  
PR 13-APR-2001; 2001US-0283674P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.

PR 12-SEP-2001; 2001US-0318770P.  
 PR 27-SEP-2001; 2001US-0325430P.  
 PR 27-SEP-2001; 2001US-0325681P.  
 PR 18-OCT-2001; 2001US-0330380P.  
 PR 31-OCT-2001; 2001US-0335012P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338092P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.

XX (PAD1/) PADIGARU M.  
 PA (SPYT/) SPYTEK K A.  
 PA (SHEN/) SHENOV S G.  
 PA (TAUP/) TAUPIER R J.  
 PA (PENA/) PENA C E A.  
 PA (LILL/) LI L.  
 PA (ZERNH/) ZERNHUSEN B D.  
 PA (GUSE/) GUSEV V Y.  
 PA (JIMW/) JI M.  
 PA (GORM/) GORMAN L.  
 PA (MILL/) MILLER C E.  
 PA (KERU/) KERUDA R.  
 PA (PATT/) PATTURAJAN M.  
 PA (GANG/) GANGOLLI E A.  
 PA (VERN/) VERNET C A M.  
 PA (GUOX/) GUO X S.  
 PA (TCHER/) TCHERNSEV V T.  
 PA (FERN/) FERNANDES E R.  
 PA (CASM/) CASMAN S J.  
 PA (MALY/) MALYANKAR U M.  
 PA (GERL/) GERLACH V.  
 PA (LITV/) LITV Y.  
 PA (ANDER/) ANDERSON D W.  
 PA (SPAD/) SPADERNA S K.  
 PA (CATT/) CATTERTON E.  
 PA (LEIT/) LEITE M W.  
 PA (ZHON/) ZHONG H.  
 PA (ALSO/) ALSOBROOK J P.  
 PA (LEPL/) LEPLLEY D M.  
 PA (RIEG/) RIEGER D K.  
 PA (BURG/) BURGESS C E.

XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;  
 PI Zernusen BD, Gusev VY, Ji M, Gorman L, Miller CE, Keruda R;  
 PI Patturajan M, Gangolli EA, Vernet CM, Guo XS, Tchernsev VT;  
 PI Fernandes ER, Casman SJ, Malynkar UM, Gerlach V, Liu Y;  
 PI Andbrook DV, Spaderna SK, Catterton E, Leite MW, Zhong H;  
 PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
 XX WPI: 2004-225693/21.  
 DR P-PSDB: ADN62079.

PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,  
 PT infection or obesity, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.

XX Claim 17; SEQ ID NO 347; 786bp; English.

XX The invention relates to an isolated polypeptide (designated NOVX, or  
 CC NOVX-NOV12) comprising a sequence selected from 178 fully defined amino  
 CC acid sequences (and their mature forms, variants and fragments). Also  
 CC included are an isolated nucleic acid molecule encoding NOVX, a vector  
 CC comprising the nucleic acid, a cell comprising the vector, methods for  
 CC determining the presence or amount of the polypeptide or the nucleic acid  
 CC molecule in a sample, methods for determining the presence of or  
 CC predisposition to a disease associated with altered levels of expression

CC of the above polypeptide or nucleic acid molecule in a first mammalian  
 CC subject, a method for identifying an agent that binds to the above  
 CC polypeptide, a method for identifying a potential therapeutic agent for  
 CC use in the treatment of a pathology that is related to aberrant  
 CC expression or physiological interactions of the polypeptide, a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide and a method for modulating  
 CC the activity of the polypeptide cited above. The composition and methods  
 CC are useful for diagnosing, preventing or treating diseases such as  
 CC diabetes, obesity, infectious diseases, anorexia, cancer-associated  
 CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or  
 CC Parkinson's disease, immune disorders, haematopoietic disorders,  
 CC dyslipidaemias, and other chronic diseases. These may also be used in  
 CC chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The polypeptides are also useful as vaccines. The  
 CC present sequence encodes a NOVX protein of the invention.

XX SQ Sequence 2820 BP; 655 A; 879 C; 654 G; 632 T; 0 U; 0 Other;

Query Match 45.6%; Score 1840.6; DB 12; Length 2820;  
 Best Local Similarity 81.4%; Pred. No. 0;  
 Matches 2367; Conservative 0; Mismatches 134; Indels 406; Gaps 4;

QY 406 CATGAGGTGACACATGGGACTACCAAGATGATGACCAATTGCTGAGGAGAG 465  
 DB 249 CATCCAGGTGACACATGGGACTACCAAGATGATGACCAATTGCTGAGGAGAG 308  
 QY 466 GATGTACGTGCTGATTTTGAACAACCTGCTCTGACTGCGGAAATGCAACGTTGCT 525  
 DB 309 GATGTACGTGCTGATTTTGAACAACCTGCTCTGACTGCGGAAATGCAACGTTGCT 368  
 QY 526 GGTGCTTTTGAATTCAGACCGGTGGGCTTCCGGCTCCGACAGGTGCTTGCACGGAAG 585  
 DB 369 GGTGCTTTTGAATTCAGACCGGTGGGCTTCCGGCTCCGACAGGTGCTTGCACGGAAG 428  
 QY 586 TCAGGAATTTGGAAATGGCTCTAGCCGAGAGATCGTGTGTGCGGCGCAAGGTGGA 645  
 DB 429 TCAGGAATTTGGAAATGGCTCTAGCCGAGAGATCGTGTGTGCGGCGCAAGGTGGA 488  
 QY 646 CTCTACGAGGAAATGTTCTCTCAAGCTTCTTCTCCCGTTAGAGAGATCAGCGGAG 705  
 DB 489 CTCTACGAGGAAATGTTCTCTCAAGCTTCTTCTCCCGTTAGAGAGATCAGCGGAG 548  
 QY 706 AAGGAGACAGTGTCAACAGATTCTTCCAGGAGTGGCCAGACTCCCAAGCTTCGGTG 765  
 DB 549 AAGGAGACAGTGTCAACAGATTCTTCCAGGAGTGGCCAGACTCCCAAGCTTCGGTG 608  
 QY 766 ACGGAGATCATGTCCCGACAGAAAGCTGTGTTCAATTCATTCAGGTTTGGATGACCTG 825  
 DB 609 ACGGAGATCATGTCCCGACAGAAAGCTGTGTTCAATTCATTCAGGTTTGGATGACCTG 668  
 QY 826 GACTCTGTCTCTCAACATGACCAAAAGCTGTGCAAAAGCTGGGTGAGAACAGCTTCG 885  
 DB 669 GACTCTGTCTCTCAACATGACCAAAAGCTGTGCAAAAGCTGGGTGAGAACAGCTTCG 728  
 QY 886 TTCACTCTCATACGACAGTCTCTCTGAGAAAGTCTGCTCTCTGAGTCTTCTGATGTC 945  
 DB 729 TTCACTCTCATACGACAGTCTCTCTGAGAAAGTCTGCTCTCTGAGTCTTCTGATGTC 788  
 QY 946 ACCGTCAGAGCGTGGGACAGAGAGAGTCAAGTCAAGTGTGCTGCTCCCGTTACTCG 1005  
 DB 789 ACCGTCAGAGCGTGGGACAGAGAGAGTCAAGTGTGCTGCTCCCGTTACTCG 811  
 QY 1006 TTAGTTAGAGGATCTCCGGGAAACAAAGATTCACCTTGTCTGAGCGCGGATTGGT 1065  
 DB 812 ----- 811  
 QY 1066 GAGCATCAGAAAGACAAAGGTTGCTGCGATCATGAAACCGTGAAGCTCTGACACAG 1125  
 DB 812 -----GCTTGTGCGGATCATGAAACCGTGAAGCTCTGACACAG 852  
 QY 1126 TGCCAGGTGCGCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGTGGTG 1185  
 DB 1126 -----

Db 853 TCCAGCTGCCGCGC----- 868  
Qy 1186 GGGGAGAGGTGGCCCCCTTCAACCAAGCTCAGAGGCTGACGCGCTTTTGTGTTT 1245  
Db 869 ----- 868  
Qy 1246 CATCAGCTCACCCCTCGAGGCGTGTGTCGCGCTGTCTCAATCTGGAGAAAAGTGTGTC 1305  
Db 869 ----- 868  
Qy 1306 CTGAAGCCCTTCTGCCGTATGCTGTGAGGGAGTGTGGAATAGGAATCAGTGTTTGAT 1365  
Db 869 -----TCTGCCGTATGCTGTGAGGGAGTGTGGAATAGGAATCAGTGTTTGATC 918  
Qy 1366 GGTGACGACTCATGATGATCAAGGAATCGGGGAGTCTGAGCTCCGCTCTGTGTTTCAGATG 1425  
Db 919 GGTGACGACTCATGATGATGAGGATCTCGCCAGTCTGAGCTCCGCTCTGTGTTTCAGATG 978  
Qy 1426 AACATCTTCTCCAGACAGCCATGTGAGAGTACTACACCTTCTTCACTCAGTCTC 1485  
Db 979 AACATCTTCTCCAGACAGCCATGTGAGAGTACTACACCTTCTTCACTCAGTCTC 1038  
Qy 1486 CAGGACTTCTGTGCGCCCTTGTACTACTGTGTAAGGGGCTGTGAAATAGAGCAGCTCTC 1545  
Db 1039 CAGGACTTCTGTGCGCCCTTGTACTACTGTGTAAGGGGCTGTGAAATAGAGCAGCTCTC 1098  
Qy 1546 TGCCCTCTGTAGCTGTGAGAGACAAAGAGTCTATGAGCTTAACAGAGGCTTCCAT 1605  
Db 1099 TGCCCTCTGTAGCTGTGAGAGACAAAGAGTCTATGAGCTTAACAGAGGCTTCCAT 1158  
Qy 1606 ATCCACTGCTTTTGGATGAAGGCTTCTTGTGCTGTGAGGCAAGCTTAAGAGAG 1665  
Db 1159 ATCCACTGCTTTTGGATGAAGGCTTCTTGTGCTGTGAGGCAAGCTTAAGAGAG 1218  
Qy 1666 CCATGAGAGTCTCTGTGCGCTGTGCTGCCCTTCCCTGCGGGGTGAAGCAAGCTTGTGAC 1725  
Db 1219 CCATGAGAGTCTCTGTGCGCTGTGCTGCCCTTCCCTGCGGGGTGAAGCAAGCTTGTGAC 1278  
Qy 1726 TGGGCTCTCTGTGGGTGAGAGGCTTAATGCAACCCAGAGAGACCCCTGAGCGCC 1785  
Db 1279 TGGGCTCTCTGTGGGTGAGAGGCTTAATGCAACCCAGAGAGACCCCTGAGCGCC 1338  
Qy 1786 TTCCACTGCTTTTGGATGAAGCTCAAGACAAAGTGTGCTGTGCAATTAACAGCTTC 1845  
Db 1339 TTCCACTGCTTTTGGATGAAGCTCAAGACAAAGTGTGCTGTGCAATTAACAGCTTC 1398  
Qy 1846 CAAGAAAGTGTGCTTCCGATTAACAGAACTGTGACTTGTCTTCTGTGCTC 1905  
Db 1399 CAAGAAAGTGTGCTTCCGATTAACAGAACTGTGACTTGTCTTCTGTGCTC 1458  
Qy 1906 CAGCACTGTCCGTAATTTGGGGAATTTGGGGTGAATGTCAAAGGGATCTTCCCAAGAT 1965  
Db 1459 CAGCACTGTCCGTAATTTGGGGAATTTGGGGTGAATGTCAAAGGGATCTTCCCAAGAT 1518  
Qy 1966 GAGTCCGTGAGGAGTGTCTGTGTCTCCCTTATGATGCGGGATTAAGACCTCAATTAG 2025  
Db 1519 GAGTCCGTGAGGAGTGTCTGTGTCTCCCTTATGATGCGGGATTAAGACCTCAATTAG 1578  
Qy 2026 GAGCAGTGGGAAGATTTTGTCTCTCATGCTTGGCACTCCACACCTGCGGAGCTGAC 2085  
Db 1579 GAGCAGTGGGAAGATTTTGTCTCTCATGCTTGGCACTCCACACCTGCGGAGCTGAC 1624  
Qy 2086 CTGGGACAGAGATCTGTACAGAGGCGGCATGAAGACCTGTGTGCCAAGCTGAGGAT 2145  
Db 1625 -----TGCAAGCTGAGGAT 1640  
Qy 2146 CCCACTGTCAAGATACAGACCTGTATGTTTAAAGATGACAGATTAACCTGTGTGTGAG 2205  
Db 1641 CCCACTGTCAAGATACAGACCTGTATGTTTAAAGATGACAGATTAACCTGTGTGTGAG 1700  
Qy 2206 CACTCTGTGAGATGTGTATGGCCAAACGTTAACTTAAGTCCCTCAACTTGGAGGACCC 2265  
Db 1701 CACTCTGTGAGATGTGTATGGCCAAACGTTAACTTAAGTCCCTCAACTTGGAGGACCC 1760

Qy 2266 CACTTGAAGAAAGAGATGTAAAGATGAGGCTGTGAAGCTTTAAACACCCAAATGTTTG 2325  
Db 1761 CACTTGAAGAAAGAGATGTAAAGATGAGGCTGTGAAGCTTTAAACACCCAAATGTTTG 1820  
Qy 2326 TTGAGTCTTTGAGGCTGAGATTTGTGTGATTTGAACCATGCTGTAACTGAAATCTCC 2385  
Db 1821 TTGAGTCTTTGAGGCTGAGATTTGTGTGATTTGAACCATGCTGTAACTGAAATCTCC 1880  
Qy 2386 CAAATCTTAAAGACCTTCCCAAGCTGAATCTGTGACCTGTGAGAGAAACAAAGTACA 2445  
Db 1881 CAAATCTTAAAGACCTTCCCAAGCTGAATCTGTGACCTGTGAGAGAAACAAAGTACA 1940  
Qy 2446 GACCAAGGAGTAATGCTCTCAGATGATGCTTGAAGATCTTCCAGTCCGCTCAGAA 2505  
Db 1941 GACCAAGGAGTAATGCTCTCAGATGATGCTTGAAGATCTTCCAGTCCGCTCAGAA 2000  
Qy 2506 CTGATCTGAGAGACTGTGATCAACCAAGGCTTGAAGATCTGAGCTTCAAGCTTC 2565  
Db 2001 CTGATCTGAGAGACTGTGATCAACCAAGGCTTGAAGATCTGAGCTTCAAGCTTC 2060  
Qy 2566 GTGACCAACCGGAGCTTGAACACCTGTGCTTATCCAAACAGCTGCGGAGAAAGT 2625  
Db 2061 GTGACCAACCGGAGCTTGAACACCTGTGCTTATCCAAACAGCTGCGGAGAAAGT 2120  
Qy 2626 GTAAATCTACTGTGTGATCATGAGCTTCCCACTGTAGTCTGACAGGCTGATGCTG 2685  
Db 2121 GTAAATCTACTGTGTGATCATGAGCTTCCCACTGTAGTCTGACAGGCTGATGCTG 2180  
Qy 2686 AATGATGCACTGTGACACAGGCTGTGTGTTTCTTGTGACTTGTGCGCTTAAGGTTAC 2745  
Db 2181 AATGATGCACTGTGACACAGGCTGTGTGTTTCTTGTGACTTGTGCGCTTAAGGTTAC 2240  
Qy 2746 TCAATGCTGACGACCTGAGCTTATGATGAACCTGTGAAAGCAATGGGTGAAGCTT 2805  
Db 2241 TCAATGCTGACGACCTGAGCTTATGATGAACCTGTGAAAGCAATGGGTGAAGCTT 2300  
Qy 2806 CTGTGAGAGTCAATGAGAAACATCTTGTATCTCCAGGACTGTGATGTTAAAGTGT 2865  
Db 2301 CTGTGAGAGTCAATGAGAAACATCTTGTATCTCCAGGACTGTGATGTTAAAGTGT 2360  
Qy 2866 CATCTCACCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTTGAGAGACAGACCTG 2925  
Db 2361 CATCTCACCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTTGAGAGACAGACCTG 2378  
Qy 2926 AAGAGCTGTGATCTCAGGGAATAGCCCTGTGTGACGCTGTGCGTGTGAG 2985  
Db 2379 -----CGAACAATGCTGTGTGACCTGTGCGTGTGAG 2422  
Qy 2986 GGACTGAAGCAAAAGAACAGTGTCTGTGACGAGCTCGGGTTGAAGGATGTGACTACT 3045  
Db 2423 GGACTGAAGCAAAAGAACAGTGTCTGTGACGAGCTCGGGTTGAAGGATGTGACTACT 2482  
Qy 3046 TGTGATGCTGTGAGGACTCTCTTGGCCCTTCTGTGCAACCGGACTGTGACAGTCTA 3105  
Db 2483 TGTGATGCTGTGAGGACTCTCTTGGCCCTTCTGTGCAACCGGACTGTGACAGTCTA 2542  
Qy 3106 AACCTGTGAGAAATTAATCAAGTCCCAAGAAATGAATGAAGCTGTGTGGCTTGGCC 3165  
Db 2543 AACCTGTGAGAAATTAATCAAGTCCCAAGAAATGAATGAAGCTGTGTGGCTTGGCC 2602  
Qy 3166 TGTCCAGCTTAATTAATTAATGAGGCTGTGAAATGAGCACTGCTGCAATA 3225  
Db 2603 TGTCCAGCTTAATTAATTAATGAGGCTGTGAAATGAGCACTGCTGCAATA 2662  
Qy 3226 AGGAAGCTGTGAGAGAGTGTGACGCTTCAAGCCCGGAGTGTGAATGAAGGTTGG 3285  
Db 2663 ACCAAGCTGTGAGAGAGTGTGACGCTTCAAGCCCGGAGTGTGAATGAAGGTTGG 2722  
Qy 3286 CATCTTTTGAATGAAGTGAACGACAC 3312  
Db 2723 CATCTTTTGAATGAAGTGAACGAGTAC 2749



RESULT 13  
ABK48611  
ID ABK48611 standard; cDNA, 3447 BP.  
XX AC ABK48611;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Mouse MATER cDNA.  
XX  
KW Mouse; gene; ss; contraceptive; antiinfertility; MATER;  
KW maternal antigen that embryos require; MATER null phenotype; oocyte;  
KW early embryonic survival; premature ovarian failure; POF;  
KW autoimmune infertility; gene therapy; fertility.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 9..3344  
FT /tag=a  
FT /product="Mouse MATER"  
XX  
FN MO200232955-A1.  
XX  
BD 25-APR-2002.  
XX  
XX 04-APR-2001; 2001MO-US010981.  
XX  
PR 18-OCT-2000; 2000US-0241510P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Nelson LM, Tong Z;  
XX  
DR MPI; 2002-454595/48.  
DR P-PSDB; AAU79525.  
XX  
XX  
PT New isolated human Maternal Antigen That Embryos Require protein and  
PT polynucleotide, useful in diagnosing or treating fertility or reduced  
PT fertility, or as a contraceptive.  
XX  
XX  
PS Example 5; Page 72-74; 93pp; English.  
XX  
XX The invention discloses an isolated human MATER (Maternal Antigen That  
XX Embryos Require) protein, which can complement a MATER null phenotype in  
XX which zygotes arising from the oocyte do not progress beyond the two-cell  
XX stage. MATER is required for early embryonic survival and abnormal levels  
XX of the protein can lead to premature ovarian failure (POF) and can be  
XX caused by under or over expression of MATER or an autoimmune response to  
XX MATER. MATER is a single-copy maternal effect gene. The MATER protein and  
XX polynucleotide, by gene therapy, are useful in diagnosing or treating  
XX fertility and reduced fertility. In particular, the MATER protein is  
XX useful as a contraceptive agent, or for influencing (either inhibiting or  
XX enhancing) fertility and can be used to detect a predisposition to  
XX infertility or reduced fertility, or for presymptomatic screening of an  
XX individual for infertility/reduced fertility. The protein and  
XX polynucleotide are also useful for detecting an excess or deficiency, or  
XX genetic mutation, of the MATER protein in a mammalian subject (e.g. a  
XX human or a mouse) or for screening for a compound useful in influencing  
XX MATER-mediated fertility. The sequence presented is the mouse MATER cDNA  
XX  
SQ Sequence 3447 BP; 1000 A; 783 C; 843 G; 821 T; 0 U; 0 Other;  
Query Match 31.7%; Score 1278; DB 6; Length 3447;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 908; Indels 18; Gaps 6;  
QY 498 TGACTGCGCCGGAATGCAACGTTGGCTGCTTTGATTGACGCGGTGGGCTTCCG 557  
DB 515 TGACGACCCAGAGATGAATATTATGTGATGCTTTTAAACATACAGAAACCTTCCA 574  
QY 558 GCTTCGACGCGTGTTCGACGGAAGTCAAGATTGGGAATGGCTTACGCAAG 617

DB 575 GCTTCACGACATTATCTACATGGAAGACAGAGATTGGGAAGTCAGTTTGGCCAGAG 634  
QY 618 GATGCTGTGTGCTGCGGCGCAAGGTGACCTTACAGGGAATGTTCTCTACCTTTCTT 677  
DB 635 TATTGTTCTTGGCTGGGCAAGGGTAACTCTTCCAAAATG---TCTTTTGTATCTT 691  
QY 678 CTTCCCGCTTGAAGATGACGCGGAAGAAAGAGAGAGTGTCAAGAGTTCACTTCAG 737  
DB 692 CTCTCTGTAGAAATTAAGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751  
QY 738 GAAGTGGCAGACTCCGAGCTCCGAGTCCGAGATCATGTCCGAGCAGAAAGCTGTT 797  
DB 752 GAAGTGTCAACATCTCGGAGATCTAGTGACAAATATCATGTCCAGAACAGAAAGCTTT 811  
QY 798 GTTCATCATTTAGAGGTTTGCATGACCTGGGCTCTGTCTC---AACATGACACAAAGCT 854  
DB 812 GTTTGTATAGATGAGCTTGATGATGATGATGATGATGATGATGATGATGATGATG 871  
QY 855 CTGCAAAAGACTGGGCTGAGAAAGAGAGCTCCGCTTACCTTATAGCACTGTGAGGAA 914  
DB 872 ATCCAGAGACTGGAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 931  
QY 915 GATCGCTCTCCGAGTCTCTCTGATGCTGACCGGTGAGAGAGAGAGAGAGAGAGAG 974  
DB 932 GCTCTCTTACTCTGATCTTCTTCTCATCATTAACACAGAAACAGAGCTTAAAGAACT 991  
QY 975 CAAGTCAGAGGTGTGTCTCCCGCTTACCTGTGATGAGAAATCTCCGGGAGAACAAAG 1034  
DB 992 CAAGTCATATGTTGTGTCTCCCGCTTATATATCTGTGAGAGAGAGAGAGAGAG 1051  
QY 1035 AATCATCTTCTCTGAG 1094  
DB 1052 ATCTAGCTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1111  
QY 1095 GATCATGAAACAACCGTGAAGTGTCTGACAGAGTGCAGAGTGCAGAGTGCAGTCTCAT 1154  
DB 1112 TCTGATGAAATATACAGAGCTGTTTGAACATGACAGAGAGAGAGAGAGAGAGAG 1171  
QY 1155 CTGCGTGGCCCTGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1214  
DB 1172 CTGAGAGGCTCTACAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1231  
QY 1215 GCTCACAGGCTGACAGCGCTTTTGTGTTTCACTACAGCTCACCCCTGAGGCGTGTCCG 1274  
DB 1232 TCTCACAGGCTTGTATGACAGAGTGTGTTTCACTACAGCTCACCCCTGAGGCGTGTCCG 1291  
QY 1275 GCGCTGTCTCAATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1334  
DB 1292 GAGGCTCTCTGAG 1351  
QY 1335 GGAAGTGTGAAATGAGAGAGTGTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1394  
DB 1352 AGAGTGTGAGACATGAGAGTGTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411  
QY 1395 GGAAGTGTGAGTGTGCTGTCTGTTTCACTAGAGAGAGAGAGAGAGAGAGAGAGAG 1451  
DB 1412 GGAAGTGTGAGATCTTGGCCCTCTTCACTAGAGAGAGAGAGAGAGAGAGAGAGAG 1471  
QY 1452 TGAAGAGTACACCTTCTTCCACCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAG 1511  
DB 1472 TGAAGAGTGTATGTTTCTTCCACCTGACCTGACAGAGAGAGAGAGAGAGAGAGAG 1531  
QY 1512 CGTGTAGAGAGGCTGGAATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1571  
DB 1532 TGTTTGAAGAGGCTGGA---GGAATGGAATGAGAGAGAGAGAGAGAGAGAGAGAG 1588  
QY 1572 GAGGTCAATGAGAGCTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1631  
DB 1589 GAGCATATGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645  
QY 1632 CTGTGTGGGCTGTGAG 1691

Db	1646	CTTATTTGGCCATCAAGAACAGAGAAATCTTGAAAGCTGTGAGAGTTCTGTTTGAATATCC	1705
Oy	1692	CGTTCCTCCGCGGGGTGAAGACAGAAAGCTTCTGCACTGGGGTCTCTCTGTGGGTCAAGACC	1751
Db	1706	CGTGAATTCACACTGTGTGACAGAAAGCTCAACAACCTGGGTCTCTCTGAATGCTCAAGCAAGT	1765
Oy	1752	TAAATGCCACCAACCCAGAGACACCTGTGAGCGCTTCCACTGTCTTTTCGAGACTCAAGA	1811
Db	1766	CAATGGCACCAAGCCCAATGACACCCCTGATGCTCTTCTAATGTCTAATTTGATGCTCAAGA	1825
Oy	1812	CAAAAGATTGTTCGCTTGGCATTTAAACAGCTTCCAGAAAGTGTGGCTTCGATTAACCA	1871
Db	1826	TGAAGAGTTTGTGTGGCGGGGCTCAACAGCTTCCAAACAAAGTGTGGCTGCTGATTAACCA	1885
Oy	1872	GAACCTGGAATTGATGACATCTTCTTGCTCCAGACACTGTCCGATATTTTCGAAAT	1931
Db	1886	GAAGATGACATGGAAGTCTCTCTCTACTGTCTCAAGCACTGTCAAGACTTGAAGGCAAT	1945
Oy	1932	TGCGGTGGAATGCAAGAGGATCTTCCCAAGAGATGAGTCCGCTAGAGCATGCTCGTGGT	1991
Db	1946	CCGGGTGATATACAGAGACTCTCTCTCGATGATATATCTCTCAAGCTGTGCCCTGTGT	2005
Oy	1992	CCCTCTATAGATGCGGGA---TAAAGACCTCATTTGAGGACAGAGTGGAAATTTCTGCTC	2048
Db	2006	TACTGTCCAGAGAGACACATATGTAAAGCCCTCTCATGSAATGTGTGGGAAACTTGTGTC	2065
Oy	2049	CATGCTTGACCAACCCACCAACCTGCGGAGCTGACCTGTGGCAGACAGCATCTTGACGA	2108
Db	2066	TGTGCTTGGCAGCCTCCGGAACCTTGAAAGAGCTGACATTGGGCGCAGACATCTGAGTCA	2125
Oy	2109	GCGGCGCAATGAAGACCTCTGTGTGCCAAGCTGAGGATATCCACCTTGCAAGATATACAGCCT	2168
Db	2126	ACGGGCAATGAAGATATCTGTGCTCTGACCTGCGAATGATGCTTGCAAAATACAGAACT	2185
Oy	2169	GATGTTTGAATGACAGATTTACCCCTGGTGTGACAGACCTCTGAGAAATTCATATGC	2228
Db	2186	GACGTTTAAAGTGCAGAGGATATGTCTGTGCTGAAACATCTCTGGAAGCTCTTTTTCAG	2245
Oy	2229	CAACCGTAACTTAAGATCCCTCAACTTGGAGGCAACCCACTGAAGAAAGATGTAAAG	2288
Db	2246	CAATCAAAACTTTAAAGTACTCAATCTAGGGGAACCTCCATGGAAGATGATGACATGA	2305
Oy	2289	GATGGCGTGTGAAGCTTTAAAAACACCCAAATGTTTGTGAGATCTTTGAGCTGTGATGG	2348
Db	2306	GTTAGCCCTGCAAGAGGCTGAAACATCCAAAGTCTCCGTGAGACTCTGAGGTTTGAACTC	2365
Oy	2349	CTGTGAGATTGACCCATGCTGTTCATGAAAGTCTCCCAAAATCCCTTACGACCTCCCGAG	2408
Db	2366	CTGTGATTAACATCAATTTGTTATGAGATGATCTCCACGCTCTTATTTAACCAACAG	2425
Oy	2409	CCTGAAATCTCTGAGCTGTGACAGGAAACAAAGGTGACACACAGGAGTAAATGCTCTCAG	2468
Db	2426	GCTAAAGTGTCTCAAGCTGTGCCAATAAAATGAGTGGAGATTAATAAGCATATATCCCTTGG	2485
Oy	2469	TGATGCTTGAAGATCTCCCAATGCGCCCTGCAAAAGCTGATACTGAGAGCTGTGAT	2528
Db	2486	GAAATGCTTGAATGATCAATGTGTCTCATCTCAAAAGTTGATATCTGACCAACTGTGGCT	2545
Oy	2529	CACAGCCACGGGTGTGCGACAGATCTGGCTCAGCCCTGCTGACCAACCGGACTTGAACA	2588
Db	2546	CACACTGCGACGCTGCCACTTCTGTGTCTCAAGCCCTTTTCAGCAACCAAGAACTTGACCA	2605
Oy	2589	CCTGTGCTTATCAACACACGCTGTGGGAAAGAAAGGTAAATCTATCTGTCTCATCCAT	2648
Db	2666	GAGGAATCAAGAAATGTGCTCTCCAGCGCTGATATCAATGATCACTGCAACATTTGATATGA	2725
Oy	2709	TGCGTGTGTTTCTTGTGCACTTGGCTTATGAGGTAACTATGCTGTGACGCACTGAGCT	2768
Db	2726	TGCTTATGCTTCTCGGCAATGATGACTTGTCAACCAACCAAGCTGACCCACCTGAGCT	2785

OY		2769	TAGCATGAACCTGTGGAAAGCAATGGCGGAAACTTCTGTGGCAGGTCATGAGAACC	2828
Dd		2786	GACCATGAACCCCTGTAAGGAGTGEGTCATGTAACCTTCTGTGTAAAGCTTTAAGAACCC	2845
OY		2829	ATTCTGTCAATCCAGAGACCTGAGATTGGTAAAGTCAATCTCACGCCCGGCTGTGA	2888
Dd		2846	TACTTGTTAACCTTCAAAGACTGAAACTATGTGACTGTGCACACTCACAAGAACTGTGCGA	2905
OY		2889	GAGTCTGTCTCTGTGTGATCTTGAGAGACAGACACTGAAAGCCTGGATCTCACGGACAA	2948
Dd		2906	GGACCTGTGCCGTGTATGATCACMACAACCAAGCACTTAAAAAGTTTGAATCTTGTGTAACAA	2965
OY		2949	TGCCCTGGGTGACGATGGGGGTTTGCTGCAGCTGTGGCAGAGGACTGAAGCAAAAAGACAGTG	3008
Dd		2966	CGCCCTGGGTATCAAAAGAGATCAATAACCTGTGTGAGAGGACTGAAGCAAGTAGACACTC	3025
OY		3009	TCTGACAGACATCGCGGGTTGAAAGGCAATGTGACATGACTTGTGATTTGCTGTGAGGCACTTC	3068
Dd		3026	CCTGAGGAGAGACTTGGGGTGGGGGCGATGTAAGTTGACTTCCAATTGCTGTGAGGCAATTGTC	3085
OY		3069	CTTGGGCCCTTCTCTGCAACCGGCAATCTGACCAAGCTTAAACCTGGTGCAGAAATACTTTCAG	3128
Dd		3146	TACATCGGGAGATGTGTGAAGCTGTGTCTGTGCGCTTCCATATGCCCTGTCTCTAACTGGGGAT	3205
OY		3189	AATTGGGCTGTGAAATATGGCAGTACCCTGTGCAAATAAAGAAAGCTGTGAGAGATGCA	3248
Dd		3206	AATTGGCCCTGTGGAAGAGAGAGATCTATGTGCCCGGATGAGAAGACAGCTGGAGGAATGTGA	3265
OY		3249	GCTACTCAAGCCCCGAGTCGTAATTGACGGTAGTGGCAATCTTTTATGAGAATGACCG	3308
Dd		3266	GTTTGTCAAGCCCACGATGTGATTGATGGTATGTGATGTAAGTGAATGAAATGACCG	3325
OY		3309	ACAC 3312	
Dd		3326	AAAC 3329	
RESULT 14				
ID	AAD49001	standard; cDNA; 3447 BP.		
XX	AC	AAD49001;		
XX	DT	07-MAR-2003 (first entry)		
DE	XX			
XX	DE	Mouse MATER cDNA.		
KM	XX	Mouse; MATER protein; infertility; fertility; contraceptive agent;		
OS	XX	gene therapy; gene; ss.		
Mus sp.	XX			
XX	Key	Location/Qualifiers		
XX	Key	9..3344		
XX	Key	/tag= a		
XX	Key	/product= "Mouse MATER protein"		
XX	Key	/transl_except= (pos:648..650, aa:Thr)		
XX	Key	/transl_except= (pos:714..716, aa:Thr)		
XX	Key	/transl_except= (pos:768..770, aa:Thr)		
XX	Key	/transl_except= (pos:887..884, aa:Thr)		
XX	Key	/transl_except= (pos:1359..1361, aa:Thr)		
XX	Key	/transl_except= (pos:1554..1556, aa:Thr)		
XX	Key	/transl_except= (pos:1740..1742, aa:Thr)		
XX	Key	/transl_except= (pos:1869..1871, aa:Thr)		
XX	Key	/transl_except= (pos:2046..2048, aa:Thr)		
XX	Key	/transl_except= (pos:2049..2051, aa:Thr)		
XX	Key	/transl_except= (pos:2229..2231, aa:Thr)		
XX	Key	/transl_except= (pos:3216..3218, aa:Thr)		

```
FT      /transl_except= (pos:3300..3302, aa:Thr)
FT      /transl_except= (pos:3330..3332, aa:Thr)
FT      /transl_except= (pos:3333..3335, aa:Thr)
FT      /note= "No start and stop codon"
FT      /partial
XX      WO200281492-A1.
XX      17-OCT-2002.
XX      29-MAR-2002; 2002WO-US009776.
XX      04-APR-2001; 2001WO-US010981.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      Nelson LM, Tong Z;
XX      MPI; 2003-058494/05.
XX      P-PSDB; AAE31748.
XX      New isolated variant MATER proteins and nucleic acids, useful for
XX      PT diagnosing, prognosing and treating infertility and reduced fertility,
XX      PT and as contraceptive agents.
XX      PS Example 22; Page 83-85; 110pp; English.
XX      CC The present invention relates to novel MATER proteins and polynucleotides
XX      CC encoding such proteins. The MATER proteins are essential to fertility.
XX      CC Sequences of the invention are useful for diagnosing, prognosing and
XX      CC treating infertility, reduced fertility and as contraceptive agents. They
XX      CC are also useful in gene therapy. The method is useful for detecting a
XX      CC predisposition to or pre-symptomatic screening of an individual for
XX      CC infertility or reduced fertility. The present sequence is mouse MATER
XX      CC cDNA
SQ      Sequence 3447 BP; 1000 A; 783 C; 843 G; 821 T; 0 U; 0 Other;
Query Match      31.7%; Score 1278; DB 8; Length 3447;
Beat Local Similarity 67.2%; Pred. No. 0;
Matches 1898; Conservative 0; Mismatches 908; Indels 18; Gaps 6;
QY      498 TGACGCGCGGAAAGCAACGTTGGCTGCTGTTGATTGACACCGGTGGGCTTCCG 557
DB      515 TGACGCGCGGAAAGCAACGTTGGCTGCTGTTGATTGACACCGGTGGGCTTCCG 574
QY      558 GCGTCGACGCGGTGTTGCGACGGAAGTCAGAAATGGGAATGCGCTCTAGCCAGAAG 617
DB      575 GCGTCACACCATTAATCTCAAGAGAGACGAGAGTTGGGAAGTCAGCTTGGCCAGAAG 634
QY      618 GATCGTCTGTGCTGCGGCGGAGGTGAGCTCTACCAAGGAATGTTCTCTACGTTCTT 677
DB      635 TATTTGTTCTGGCTGGGCGACAGGTAACTTCTCCAAAAAATG---TCCTTGTGATCTT 691
QY      678 CTTCCCGCTTGAAGATGACGAGCGAAGAGAGAGCGTGTACAGAGTTATCTCCAG 737
DB      692 CTTCTCTGTAGAGAAATTAAGTGAACAGAGAAAGCGTTGGACAGCTGATTTGCTAA 751
QY      738 GAGAGTGGCAGACTCCAGAGCTCCGATGACGAGATCATGTCCCAACGAGAAAGCTGTT 797
DB      752 GGAAGTGTCCAGACTCTCGGATCTAGTGAACAAAGTATGATCTCCCAACGAGAAAGCTGTT 811
QY      798 GTTCATCATGACGCTTTCATGACCTTGGGCTCTGTCTC---AACATGACACAAAGCT 854
DB      812 GTTTCTCATATGAGCTTGGATGATATGAGACTCTGTCTCCAAATGATGATATGACT 871
QY      855 CTGCAAGACGCGGTGAGAGACGCTCCGTTACACCTTCATACCGCAGTCTCTAGGAA 914
DB      872 ATCCAGAGCTGGAAGATGAACAGCCCATATATACCTCTGATGTAACAGCTCTTAGGAA 931
QY      915 GGTCTGTGCTCCCTGAGTCTCTCTGATGTCACCGTCAGAGAGCTGGGACAGAGAACT 974
DB      932 GGCCTCTTACTACTAGTCTCTTCTCATATTAACACAGCAAGACAGGCTTAGAAAACT 991
```

```
QY      975 CAAGTCAGAGGTCGCTGCTCCCGTTACCTGTAGTAGAGAAATCTCCGGGACAAAG 1034
DB      992 CAAGTCAGAGGTCGCTGCTCCCGTTACCTGTAGTAGAGAAATCTCTGATCAAGAG 1051
QY      1035 AATCCACTTGTCTCTTGAACGCGGAGTGTGAGCATGAGAGACAAAGGTTGCTGTC 1094
DB      1052 ATCTGAGCTGTGCTCGAGAAACATCTCCATAGCTGATAGAAATACAAAGTCTTCATTC 1111
QY      1095 GATCATGAAACAACGCTGAGCTGTGACACAGTGCAGGTGCCGCTGCTCTCAT 1154
DB      1112 TCTGATAGAAATATCACAGCTGTTTGACCAATGCGAGGCCCTCTGTGTGCTCTGCT 1171
QY      1155 CTGCGTGGCCCTGACGCTGACAGAGCTGTGGGAGAGCGTGGCCCTTCAACCAAC 1214
DB      1172 CTGTAGGCTCTACAGCTACAGAAACCTGGGAAAGATGACCTTACCTGCTGACAC 1231
QY      1215 GCTCACAGGCTGACGCGCTTGTGTGTTTATCATGACTCACCCCTCGAGGCGTGTGCG 1274
DB      1232 TCTCACGCGTTGTATGTCACAGTGTGTTTTCACAGCTCACCTTGAAGGCTTCCCA 1291
QY      1275 GCGTGTCTCAATCTGAGAGAAAGATTGCTTGAAGGCTTTCGCGTATGCTGTGGA 1334
DB      1292 GAGCGCTCTCAGTCAGAGAAACAGATTAATCTAGTGGGTTGTGTCATGATGCGAGCTGA 1351
QY      1335 GGAAGTGTGGAATGAGAGTCAAGTGTGATGATGACACCTTCATGCTTCAAGACTGCG 1394
DB      1352 AGGAGTGTGACCAATGAGAGTGTGTTATGATATATACCTGAAACCTATAGCTTAA 1411
QY      1395 GGAAGTGTGAGCTCGGTGCTGTTTACATGAAACCTTCTCCAGACAGC---ACTG 1451
DB      1412 GGAAGTGTGATCTTGGCCCTCTTTCACATGAACTCTTTCAGAGTGTGGCCACAAAG 1471
QY      1452 TGAGAGTACTACACTTCTTCCACCTGACGTCTCCAGAGCTTGTGCGGCTGTACTA 1511
DB      1472 TGAGCAGTGTATGTTTCTCCACACTCAGCTGACGAGATTTCTTGTGCTTATATTA 1531
QY      1512 CGTGTTAAGAGGCTGGAATATGAGCCAGCTCTGCGCTCTGTATGATGAGACAA 1571
DB      1532 TGTTTTAAAGGGCTGGA---GGAATGGAATCAGATTTTGTGCTTATGAAACCAAG 1588
QY      1572 GAGGTCAATGAGAGCTTAAACAGGCGCTTCATATCACTCGCTTGTGATGAGCGTTT 1631
DB      1589 GAGCATCAATGAGGTTGAAGAACTGACGACACTGCGCTC---CTCGGAGTGAACGTTT 1645
QY      1632 CTTGTTGGCTCTGAGCGAAGAGCTGAAGAGGCACTGAGAGTCTGTGGGCTGTCC 1691
DB      1646 CTTATTTGGCTCATGAGCAAGATATCTTGAAGCTCTGAGAGTCTGTTGATATACC 1705
QY      1692 CGTTCCTCGGAGGTGAAGCAAGAGCTTGTGACTGTCTCTGTGGGTCAAGACCC 1751
DB      1706 CGTGAATTCACATGTTGAGCAAGAGCTTCCAACTGAGGCTCTCTGATAGCTCAGCAGGT 1765
QY      1752 TAATGCAACACCCAGAGACACCTGAGAGGCTTCCAGCTGCTTTCGAGACTCAGA 1811
DB      1766 CAATGCAACAGCCCAATGAGACACCTGAGAGGCTTCTATATGTCTAATTTGAGTCAAG 1825
QY      1812 CAAGAGTGTGCTTGTGCTTGAATTAACAGCTTCCAGAGAGTGTGCTTCCGATTTAAC 1871
DB      1826 TGAAGAGTTGTGTGGCGGGGCTCTCAAAACGCTTCCAAAGATGTGGCTGATTTAAC 1885
QY      1872 GAACTGGACTTATAGATCTTCTCTGCTCTCAGACGCTGTCCGATTTTGGGAAAT 1931
DB      1886 GAAAGATGACCTTGAAGGCTCTCTCTCTACTGTCTCAAGACACTGTGAGAACTGAAAGCAAT 1945
QY      1932 TGGGTGATGATCAAGAGGATCTTCCAAAGAGATGAGTCCGCTGAGAGGATGTCCTGTGGT 1991
DB      1946 CCGGATGATATCAAGAGCTCTCTGAGTATATATCTCTGAGAGTGTGCTGCTGTGT 2005
QY      1992 CCTCTATGATGCGGGA---TAAGACCTCATTTGAGAGAGAGTGTGGAAGATTTCTGTGCTC 2048
DB      2006 TACTGTCAAGAGACAAATGATTAAGCCCTCTCATGAGAGTGTGGGGAACCTTCTGTGCTC 2065
```

QY	2049	CATGCTGGGCAACCAACCCACACCTGGCGGCACTGGACCTGGGGCAGACGATCCGACAGA	2108
Db	2066	TGTGCTTGTGGCAGCCTCCGGAACTTGAAAGACCTGGACTTTGGGCGACAGACATCCGAGTCA	2125
QY	2109	GCGGGCCATYGAAAGACCCTGTGTGCGCAAGCTGAGGCATCCCACTGCGAAAGATACAGACCCCT	2168
Db	2126	ACGGGCCCATGAAAGATACTGTGCTCTGAGCTCCGAAATACAGTCTCTGAGAAATACAGAAAGCT	2185
QY	2169	GATGTTTGAATATGACACAGATTTACCCCTGGTGTGACGACCTCTGGAGAAATGTCATAGCC	2228
Db	2186	GACGTTTAAAGTGCAGAGGTAGTGTGTGCTGGCCTGAACATCTCTGAAAGCTCTTTTATAG	2245
QY	2229	CAACCCGTACCTTAAGATACCCCTCAACTTGGGAGGCAACCACTGAAAGGAAGAGATGTAAG	2288
Db	2246	CAATCTAAACCTTAAAGTACCTCAATCTAAGGAAACCTCCAAATGAAAGATGATGACATGAA	2305
QY	2289	GATGGCCGTGTAAAGCCTTAAACACCCAAATGTTTGTGTGAAGTCTTTGAAGGCTGAGATTG	2348
Db	2306	GTTAGCCTGGGAAGGCTGAACATCCAAATGTCTCCGTGAGACATCTGAGGTTGGATTTC	2365
QY	2349	CTGTGGAATTTGACCCATGCTGTATCTGAAAGATCTCCCAATTCCTTACGACCTTCCCGAC	2408
Db	2366	CTGTGAGTTAACATCATTTGGTTATAGAGATGATCTCCAGCCTTCTATTTCACACACAG	2425
QY	2409	CCTGAAATCTCTGAGCCTGGCGAGAAACAAAGTGCACAGACGAGAGTAAATGCTCTACAG	2468
Db	2426	GCTAAAGTGTCAACGCTCGGCGCAAAATTAAGTGGAGTAAAAAGCAATGATATCCCTTGG	2485
QY	2469	TGATGCTCTTGAAGAGTCTCCCAAGTGGCGCCTGTGACGAAGCTGTATCTGAGAGACTGTGGCAT	2528
Db	2486	GAAATCCCTTGAAGTACCTAATGTGTCTACTGCAAAAGTTGATATCTGACCAACTGTGGCCT	2545
QY	2529	CACAGCCACGGGTTGCCAAGTCTTGACCTCTGACGAACTGGACCTTGAACA	2588
Db	2546	CACACCTGCGAGCTGCGACACTTCTGTGTCTGACCCCTTTTCAGCAACGAACTTGAACA	2605
QY	2589	CCTGTGCTATTCACAAACAACGCTTGGGGAACGAAGGTAAATCTACTGTGTGCATCCAT	2648
Db	2606	CCTGTGCTGTGCAAAACAACGCTTGGGGAACGAAAGATGCAACAGCTGTGTCAAGTTCTT	2665
QY	2649	GAGGCTTCCCACTGTAGTCTGCGAGAGGCTGATGTGAATCAGTGGCCACTTGAACAACGCG	2708
Db	2666	GAGGAATTCAGAAATGTGCTCTTCCAGCGCTGATCTGAAATCACTGCAACATTTGATAGTA	2725
QY	2709	TGGCTGTGGTTTCTTGACACTTGGCTTATGGGTAACTCATGGCTGACGACCTGAGCTT	2768
Db	2726	TGCTTATGGCTTCCGGCAATGAGACTTGCAAACAACAAGCTGACCCACCTGAGGCTT	2785
QY	2769	TAGCATGAACCTCTGTGGAAGACAAATGCGGTGAAGCTTGTGCGAGGTTCATGAGAGAAC	2828
Db	2786	GACCATGAAACCCCGTAGGGAGATGGTGCAAATAAAGGTACTGTGTAAAGCTTTTAAAGAACT	2845
QY	2829	AATCTGTATCTCCAGACCTGAGATGGTGTAAAGTGCATCTCACCCGCGGTGCTGGA	2888
Db	2846	TACTTGTTAACCTTCAAGAACCTGGAACTAAGTGACTGCCAATCACAACGAACCTGTGCGA	2905
QY	2889	GAGTCTGTCTGTGTGATCTCGAGAGGACGACACTGAAAGACCTGTGATCTCACGACAA	2948
Db	2906	GGACCTGGCTGTATGATCACAAACAACAACACTTAAAGTTTGGATCTTGGTAAACA	2965
QY	2949	TGCGCTGGGTGACGGTGGGGTTGCTGCGCTGTGCGAGGACTGAAGCAAAAGAACAGTGT	3008
Db	2966	CGCCCTGGGTGACAAAGGAGTCATTAACCTGTGTGAGGGACTGAAGCAAAATGACAGCTC	3025
QY	3009	TCTACACGAGACTCGGGTGTGAAGGACATGGAACATACTTCTGATGTCTGTAGGGAACCTCTC	3068
Db	3026	CCTAGAGGACCTTGGGTGTGGGCGCATGTAAAGTTGACTTCAATTTGCTGTGGCATTTGTC	3085
QY	3069	CTTGGCCCTTCTCTGCAACCGGCAATCTACCAAGTCTTAAACCTGGTGCAAGTAACTTTCAG	3128
Db	3086	ATTGGCCATCTCTTGCACAACCTCAACCTGAACAGCTTAAACCTGGTGAAGAAATGACTTCAG	3145
QY	3129	TCCCAAGGAATGATGAAGCTGTGTTGGGCTTTGCTGTGCCACGCTTAACTTACAGAT	3188

Db	3146	TACATCGGAGATGTTGAAGCTGTGCTTCGCGTTCCAAATGCCCTGCTCTTAACCTGGGGAT	320
Qy	3189	AATGGGCGTGTGAAATGGCAGTACCCCTGTGCAAAATGAAGAGCTGTGGAGAAATGCA	324
Db	3206	AATGGCGCTGTGGAAGCAGAGACTATATGCCGAGTGAAGACAGCTGGAGGAAGTTGA	326
Qy	3249	GCTACTCAAGCCCCGAGTCGTAAATGACGGTAGTTGGCATTTTGTATGAATGACCG	330
Db	3266	GTTCGTCAAGCCCAACGTGTGATTGAATGTGATGTATGCTAAGTGAAGAAATGACCG	332
Qy	3309	ACAC 3312	
Db	3326	AAAC 3329	
RESULT 15			
ID	ABK48609		
XX	ABK48609	standard; CDNA, 1157 BP.	
AC	ABK48609;		
XX			
DT	13-AUG-2002	(first entry)	
XX			
DE	Human MATER cDNA fragment #1.		
XX			
KW	Human; gene; ss; contraceptive; antiinfertility; MATER;		
KW	Maternal antigen that embryos require; MATER null phenotype; oocyte;		
KW	early embryonic survival; premature ovarian failure; POF;		
KW	autoimmune infertility; chromosome 19; gene therapy; fertility.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1155	
FT		/tag= a	
FT		/product= "Human MATER"	
FT		/partial	
FT		/note= "No start or stop codon shown"	
FN	WO200232955-A1.		
PD	25-APR-2002.		
XX			
XX	04-APR-2001;	2001WO-US010981.	
PR	18-OCT-2000;	2000US-0241510P.	
XX			
PA	(USSH )	US DEPT HEALTH & HUMAN SERVICES.	
XX			
PI	Nelson LM, Tong Z;		
XX			
DR	WPI; 2002-454595/48.		
XX	P-PSDB; AAU79523.		
PT	New isolated human Maternal Antigen That Embryos Require protein and		
PT	polynucleotide, useful in diagnosing or treating fertility or reduced		
XX	fertility, or as a contraceptive.		
PS	Claim 11; Page 68-69; 93pp; English.		
XX			
CC	The invention discloses an isolated human MATER (Maternal Antigen That		
CC	Embryos Require) protein, which can complement a MATER null phenotype in		
CC	which zygotes arising from the oocyte do not progress beyond the two-cell		
CC	stage. MATER is required for early embryonic survival and abnormal levels		
CC	of the protein can lead to premature ovarian failure (POF) and can be		
CC	caused by under or over expression of MATER or an autoimmune response to		
CC	MATER. MATER is a single-copy Maternal effect gene found on chromosome		
CC	19. The MATER protein and polynucleotide, by gene therapy, are useful in		
CC	diagnosing or treating fertility and reduced fertility. In particular,		
CC	the MATER protein is useful as a contraceptive agent, or for influencing		
CC	(either inhibiting or enhancing) fertility and can be used to detect a		
CC	preldisposition to infertility or reduced fertility, or for presymptomatic		

screening of an individual for infertility/reduced fertility. The protein and polynucleotide are also useful for detecting an excess or deficiency, or genetic mutation, of the MATR protein in a mammalian subject (e.g. a human or a mouse) or for screening for a compound useful in influencing MATR-mediated fertility. The sequence presented is the human MATR cDNA fragment #1, which was isolated from a human cDNA library

Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Query Match 28.6%; Score 1152.8; DB 6; Length 1157;  
Best Local Similarity 99.8%; Pred. No. 9.3e-308;  
Matches 1154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 754 CAGGCTCCGGTACCGAGATCATGTCGCCGACGAAAGGCTTTGTCATCTTACAGCT 813
DB 1 CAAAGTCGGGAGACGAGATCATGTCGCCGACGAAAGGCTTTGTCATCTTACAGCT 60
QY 814 TTCGATGACCTGGGCTCTGCTCTCAACATGACACAAAGCTTGCAGAAAGCTGGGCTGAG 873
DB 61 TTCGATGACCTGGGCTCTGCTCTCTCTCAACATGACACAAAGCTTGCAGAAAGCTGGGCTGAG 120
QY 874 AAGCAGCCTCCGTTCAACCTCATACGAGCTGTGTGAGGAGAGTCTGCTCCCTGAGTCC 933
DB 121 AAGCAGCCTCCGTTCAACCTCATACGAGCTGTGTGAGGAGAGTCTGCTCCCTGAGTCC 180
QY 934 TTCCTGATCTGACACGCTCAAGACGTTGGGCAACAGAAAGCTCAAGTCAAGAGTCTGTCT 993
DB 181 TTCCTGATCTGACACGCTCAAGACGTTGGGCAACAGAAAGCTCAAGTCAAGAGTCTGTCT 240
QY 994 CCCCGTTACCTGTAGTGTAGAGAAATCTCCGGGGAACAAAGAAATCACTGTCTCTTGGAG 1053
DB 241 CCCCGTTACCTGTAGTGTAGAGAAATCTCCGGGGAACAAAGAAATCACTGTCTCTTGGAG 300
QY 1054 CGCGGGAATTGTGAGCATCAGAAACACAAAGGTTGCGTGCATCATGAAACAAACCTGAG 1113
DB 301 CGCGGGAATTGTGAGCATCAGAAACACAAAGGTTGCGTGCATCATGAAACAAACCTGAG 360
QY 1114 CTGCTCGACCAATGTCAGAGTGCCTCCGCTGTCTCTCATCTGCTGGGCTCTGCACTG 1173
DB 361 CTGCTCGACCAATGTCAGAGTGCCTCCGCTGTCTCTCATCTGCTGGGCTCTGCACTG 420
QY 1174 CAGGACGTGTGGGGAAGAGCGTCCGCCCTCAACCAAGCTCAACAGGCTGACAGGCC 1233
DB 421 CAGGACGTGTGGGGAAGAGCGTCCGCCCTCAACCAAGCTCAACAGGCTGACAGGCC 480
QY 1234 GCTTTTGTGTTTCATCAGCTCAACCCCTGAGGCGTGTCCGCGCTGTCTCATCTGAG 1293
DB 481 GCTTTTGTGTTTCATCAGCTCAACCCCTGAGGCGTGTCCGCGCTGTCTCATCTGAG 540
QY 1294 GAAAGATTGTCTGAGAGCGCTTCTGCGCTATGCTGTGAGGAGTGTGAATAGAAAG 1353
DB 541 GAAAGATTGTCTGAGAGCGCTTCTGCGCTATGCTGTGAGGAGTGTGAATAGAAAG 600
QY 1354 TCAGTGTGTGATGTGAGGAGCTCATGCTCAAGGACTCGGGGAGTCTGAGCTCGTGTCT 1413
DB 601 TCAGTGTGTGATGTGAGGAGCTCATGCTCAAGGACTCGGGGAGTCTGAGCTCGTGTCT 660
QY 1414 CTGTTTCAATGAAATCTCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTC 1473
DB 661 CTGTTTCAATGAAATCTCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTC 720
QY 1474 CACTTCACTCTCAGGACTTCTGTGCGCGCTTGTACTACTGTGTAGAGGCGCTGGAATTC 1533
DB 721 CACTTCACTCTCAGGACTTCTGTGCGCGCTTGTACTACTGTGTAGAGGCGCTGGAATTC 780
QY 1534 GAGCAGCTCTCTGCGCTCTGACGTTGAGAGACAAAGAGTCCATGAGGCTTAAACAG 1593
DB 781 GAGCAGCTCTCTGCGCTCTGACGTTGAGAGACAAAGAGTCCATGAGGCTTAAACAG 840
QY 1594 GAGGCTTCCATATCACTCGCTTGTGAGTGAAGCGTTCTTGTGAGCTCTGAGCGAA 1653
DB 841 GAGGCTTCCATATCACTCGCTTGTGAGTGAAGCGTTCTTGTGAGCTCTGAGCGAA 900

```

```

QY 1654 GACGTAAAGAGGCCACTGAGAGTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAG 1713
DB 901 GACGTAAAGAGGCCACTGAGAGTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAG 960
QY 1714 AAGCTTTCAGACTGGGTCTCTGTGAGTCAAGGCTTAAATGCAACCAAGGAGAGAC 1773
DB 961 AAGCTTTCAGACTGGGTCTCTGTGAGTCAAGGCTTAAATGCAACCAAGGAGAGAC 1020
QY 1774 ACCGTGAGCGCTTTCATCTGTCTTTGAGACTCAAGCAAGAGTTGTTCCCTTGGCA 1833
DB 1021 ACCGTGAGCGCTTTCATCTGTCTTTGAGACTCAAGCAAGAGTTGTTCCCTTGGCA 1080
QY 1834 TTAACAGCTTCAAGAGTGTGGCTCCGATTAAACAGAACTGGAATTGATACATCT 1893
DB 1081 TTAACAGCTTCAAGAGTGTGGCTCCGATTAAACAGAACTGGAATTGATACATCT 1140
QY 1894 TCCTTCTGCTCCAGC 1909
DB 1141 TCCTTCTGCTCCAGC 1156

```

Search completed: July 18, 2005, 16:47:42  
Job time : 2052 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 16:15:50 ; Search time 632 Seconds  
(without alignments) 10446.798 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035  
Sequence: 1 atggaagagagacaatcgct.....tcgaggggctggtctctaa 4035

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405566

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221.4	5.5	4422	US-09-388-221B-1	Sequence 1, Appli
2	221.4	5.5	4422	US-09-388-221B-1	Sequence 9, Appli
3	220.6	5.5	4200	US-09-388-221B-3	Sequence 3, Appli
4	220.6	5.5	4332	US-09-388-221B-5	Sequence 5, Appli
5	220.6	5.5	4466	US-09-388-221B-11	Sequence 11, Appli
6	183.8	4.6	1371	US-08-910-731-1	Sequence 1, Appli
7	183.8	4.6	1371	US-08-910-731-5	Sequence 5, Appli
8	182.4	4.5	1386	US-08-910-731-7	Sequence 7, Appli
9	180.8	4.5	1372	US-08-910-731-3	Sequence 3, Appli
10	175.8	4.4	1374	US-08-910-731-3	Sequence 3, Appli
11	175.8	4.4	1374	US-08-910-731-3	Sequence 3, Appli
12	61	1.5	7218	US-08-232-463-14	Sequence 14, Appli
13	61	1.5	376	US-09-023-655-395	Sequence 395, App
14	51.4	1.3	29836	US-09-949-016-15600	Sequence 15600, A
15	50.8	1.3	64309	US-09-949-016-14581	Sequence 14581, A
16	47.4	1.2	4485	US-09-949-016-3209	Sequence 3209, Ap
17	47.4	1.2	4485	US-10-014-269-1	Sequence 1, Appli
18	47.4	1.2	4485	US-10-014-269-1	Sequence 33, Appli
19	46.6	1.2	1276	US-09-177-325-2	Sequence 2, Appli
20	46.6	1.2	1276	US-09-411-812A-2	Sequence 2, Appli
21	46.6	1.2	1276	US-09-590-113-2	Sequence 358, App
22	46.4	1.1	2780	US-09-620-312D-358	Sequence 14, App
23	45	1.1	929	US-09-671-317-14	Sequence 340, App
24	44.8	1.1	1620	US-09-134-000C-340	Sequence 439, App
25	43.4	1.1	1001	US-09-671-317-439	Sequence 37, Appli
26	43.2	1.1	193303	US-09-497-855A-37	Sequence 44, Appli
27	43.2	1.1	193303	US-09-497-855A-44	

28	42.8	1.1	1449	US-09-248-796A-4693	Sequence 4693, Ap
29	42.8	1.1	767677	US-09-949-016-12147	Sequence 12147, A
30	42.8	1.1	767677	US-09-949-016-11361	Sequence 11361, A
31	42.6	1.1	693	US-09-252-991A-11050	Sequence 11050, A
32	42.6	1.1	1005	US-09-252-991A-10976	Sequence 10976, A
33	42.6	1.1	1185	US-09-252-991A-11203	Sequence 11203, A
34	42.4	1.1	150394	US-09-949-016-13042	Sequence 13042, A
35	42.2	1.0	38566	US-09-949-016-15271	Sequence 15271, A
36	42.2	1.0	38566	US-09-949-016-15272	Sequence 15272, A
37	42	1.0	3117	US-09-614-221A-275	Sequence 275, App
38	41.8	1.0	7044	US-09-949-016-14113	Sequence 14113, A
39	41.6	1.0	2223	US-08-257-073-4	Sequence 4, Appli
40	41.4	1.0	505	US-09-621-976-15639	Sequence 15639, A
41	41.2	1.0	810	US-08-742-185-101	Sequence 103, App
42	41.2	1.0	43795	US-08-742-185-101	Sequence 101, App
43	41	1.0	505	US-09-621-976-15639	Sequence 15639, A
44	41	1.0	10391	US-09-902-540-958	Sequence 958, App
45	40.4	1.0	240157	US-09-949-016-16264	Sequence 16264, A

## ALIGNMENTS

RESULT 1					
US-09-388-221B-1					
Sequence 1, Application US/09388221B					
Patent No. 6818750					
GENERAL INFORMATION:					
APPLICANT: Reed, John C.					
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation					
FILE REFERENCE: P-IJ 3650					
CURRENT APPLICATION NUMBER: US/09/388,221B					
CURRENT FILING DATE: 1999-09-01					
NUMBER OF SEQ ID NOS: 30					
SOFTWARE: PatentIn Ver. 2.0					
SEQ ID NO 1					
LENGTH: 4422					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: CDS					
LOCATION: (1)..(4422)					
US-09-388-221B-1					
Query Match					
Best Local Similarity 47.9%; Pred. No. 9.1e-55;					
Matches 953; Conservative 0; Mismatches 966; Indels 72; Gaps 8;					
QY	532	TTTGAATTCAGACCGGCTGCGCTCCGCTGCGACCGCTGCTTTCGACGGAATCAGGA	591		
DB	952	TTTGCCCAAGGCTGATATACCAAGAACCTGCAATGATCACTGACAGGGGCTGCTGA	1011		
QY	592	ATTGGGAATCGGGCTTACGAGAGATCGTGTGCTGGGCGCAAGTGAAGTCTTAC	651		
DB	1012	ATTGGGAATCGAGCACTGCGACAGAGTGAAGAGCTGGGAGAGGCCAGCTGAT	1071		
QY	652	CAGGAATGCTCTCTACGCTCTTCTCCCGTTAGAGATGACAGCGGAAGAGAG	711		
DB	1072	GGGAGCGCTTCCAGCATGCTCTTCACTTCACTGACAGAGCTGCGCCAGTCCAGGTG	1131		
QY	712	AGCATGTCACAGATTATCTCCAGGAGTGGCCAGCTCCAGGCTCCGGTACGAG	771		
DB	1132	GTGATCTCGTGGAGCTCATGGAAGAAATGGAGCACTCCGGCTCCATTAGACAG	1191		
QY	772	ATCATGTCACAGAAAGGCTGTTGTTCAATCAATGAGGTTTGAATGACCTGGCTCT	831		
DB	1192	ATTCGTGTACGACGAGAGCGGCTCTCTTATCTTCATGATGATGATGACGAGATG	1251		
QY	832	GTCTCAACAAATGACA---CAAGCTTCGAAAGCTGGCTGAGAGAGCACTTCCTTC	888		
DB	1252	GTCTTGCAGAGAGCGAGTTCGACTGTGTGACATGAGACCAAGCCAGCCGCGGAT	1311		
QY	889	ACCCTCATGCGAGTCTGCTGAGAGAGGTCTGCTCCCTGAGTCTTCTTATGCTACC	948		

```
Db      1312  GACCTGCTGGGAGATTGCTGGGGAACCTACTCTCCGAGGACATCTCTCTATCAACG 1371
Qy      949   GTCAAGACGTGGGACACAGAAAGTCAAGTCAAGGTCTGTCTCCCGTTACTCTGTA 1008
Db      1372  GCTCGGACCAAGCTCTGCAAGAACTCAATCTCTTCTTGGAGCAGGCACTGTGGGTAAAG 1431
Qy      1009  GTTAGAGAAATCTCCGGGGAAACAAGAAATCACTTGTCTCTGAGCGGGGATGGTGAG 1068
Db      1432  GTCCGGGGGTTCTTGAAGTCCAGCAGAGAAAGAAATTTCTACAGATTTTCAAGATGAA 1491
Qy      1069  CATCAGAAAGACAAAGGTTGGTGCATCAATGAACAACGTGACCTGTCGACCAAGTGC 1128
Db      1492  AGGCAAGCAATTAAGACCTTTAGTTGGTCAATCAACAAAGAGCTGTGGGCCCTGTGT 1551
Qy      1129  CAGGTGCCCCCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGACAGAGCTGTGGGG 1188
Db      1552  CTGTGCCCCGTGGGTCTCTGGCTGGCTGCACTTGCCTGATGCAAGCATGAAGCGGAAG 1611
Qy      1189  GAGAGCGTGGCCCCCTTCAACCAACGCTCAAGGCTGCAAGCGCGCTTTTGTGTTCAT 1248
Db      1612  GAAAACTCAACTGACTTCCAGACACACACACACCTCTGTCTATCAATTTACTTGCAG 1671
Qy      1249  CAGCTCACCCCTCGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAAAGATGTCTCTG 1308
Db      1672  GCTCTCAAGCTCAAGCATTGG-----GACCCAGCTC 1704
Qy      1309  AAGCGCTTCTGCGGTATGCTGTGGAGGAGGTGGAATGAGAGTCACTGTGTGATGCT 1368
Db      1705  AGAGACCTCTGCTCTGTGCTGTGAGGGCACTGTGGCAAAAACACCTTTTCACTGCA 1764
Qy      1369  GAGGACCTCATGTTCAAGGACTGTGGGAGTGTGAGCTCCGTCCTGTGTTTCACTGAC 1428
Db      1765  GATGACCTCAAGAGACATGGGTTAGATGGGCGCATCTCACTTCTTGAAGATGGGT 1824
Qy      1429  ATCTCTTCCAGACAGCCACTGTGAGAGATCTACACCTTCTTCCACTGATCTCCAG 1488
Db      1825  A---TTCTTCAAGACACCCCATCCCTGAGCTACAGCTTCACTTCACTGTTTCCAA 1881
Qy      1489  GACTTCTGTGCGCTTGTACTAGCTGTTAAGAGGCGCTGGAATATGAGCCAGCTCTGC 1548
Db      1882  GAGTCTTGTGACGACATGTCTATGTCTTGTGAGATGAAGAGGGGAGATGAACATTCT 1941
Qy      1549  CCTCTGATGTTGAGAAACAAGAGTCAATGAGCTTTAAACAGGAGGCTTCCATATC 1608
Db      1942  AATTGCATCATAGATTTGAAAAGCGCTAGAAGCATATGAAATCATGCGCTGTT--- 1998
Qy      1609  CACTGCTTTGGAATGAAGCGTTTCTTGTGGCTCTGAGCGAGAAGATGAAGAGCCA 1668
Db      1999  ---GGGGCATCAACACAGCTTCTATTTGGGCTGTAAATGATGAGGGGAGAGAGAG 2055
Qy      1669  CTGAGAGTCTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGACAGAGCTTCTGCATCG 1728
Db      2056  ATGGAACAATCTTTCATCTGCCGCGCTGTCTCAAGGGGAGAACTGATGCACTGTGCTCCG 2115
Qy      1729  GTCTCTGTGGGCTCAGACGCTTAATGCAACCAACCCAGAGACACCTGAGAGGCTTC 1788
Db      2116  TCCCTGACGCTGCTGTGAGC-----ACACTCTCTGAGTCCCTC 2157
Qy      1789  CACTGTCTTTTCAAGACTCAAGACAAAGATTTGTTCCCTTGGCACTTAAGAGCTTCCA 1848
Db      2158  CACTGCTTGTAGAGAGCTCGAGACAAAGCTTCCGACACAAAGTATGCGCCATTGGAA 2217
Qy      1849  GAAAGTGGCTTCGATTAACCAAGAACTGACTTGAATGACATCTTCTTGTGCTCCAG 1908
Db      2218  GAAATGGGCACTGTGTGTGAAACACAGATGAGCTTAAAGTGTGACATTTTGCATTAAA 2277
Qy      1909  CACTGTCCGTATTTGGCGGAAATTCGGGTGATGATCAAGAGATTTCCCAAGATGAG 1968
Db      2278  TTCAAGCCCAAGTAAAGCTTCAAGCTGATGTAGGCGCAGG-----CAGCAAGATC 2330
Qy      1969  TCCGCTGAGGATGCTGTGTGCTCTCTATGATGAGCGGAGATGAAGACCTCATTTAGAGAG 2028
```

```
Db      2331  AACATGAGCCCCACCAATGATGATCTGTTCAGGTGGGTCC-----CAGTCAAGATGCC 2385
Qy      2029  CAGTGGAAAGTTTCTGTCTCCATGCTTGGACCAACCAACCTTGGCGAGCTGGACTG 2088
Db      2386  TATTGGAGATTTCTTCTTCTCCGTCTCAAGGTCAACGAAACCTGAAGAGCTGGACTTA 2445
Qy      2089  GGCAGCAGATCTCGACAGAGCGGGCCATGAAGACCTGTGTGCAAGTCAAGGCAATCCC 2148
Db      2446  AGTGAAGATCGCTGAGCCACTCTGCACTGAAGAGTCTTTGTAAAGACCTTGAAGCGGCT 2505
Qy      2149  ACCTGCAAGATACAGACCTTGATGTT--TAGAAATCAAGATTAACCTGTGTGTGAG 2205
Db      2506  CGTCCCTCCCTGAGAGACCTCGGCTGTGCTGTGCTGTGCTCAACGCTGAGGACTGGAAG 2565
Qy      2206  CACTCTGAGAAATCGTATGAGCCCAACGTTAACCTTAATCCCTCAACTTGGAGGAGC 2265
Db      2566  GACTTGTCTTTGGGCTGAGAGCCACACAGACCTTGAACCGAGCTGAGCTTGAAT 2625
Qy      2266  CACTGAAGAGAAAGAGATGTAAGATGCGGTGAACCTTTAAACACCCAAAATGTTTG 2325
Db      2626  GTGCTCAAGATGCTGAGAGCCAAACACTTTGCCAGAGACTGAGACAGCCGAGCTGCAAG 2685
Qy      2326  TTGAGTCTTTGAGGCTGATGTCTGTGATTTGACCAATGCTGTGTACTTGAAGATCTCC 2385
Db      2686  CTACAGGCACTGCACTGTGTGAGCTGTGAGCTTCACTGATCTGTGCTCCAGGACTGAGC 2745
Qy      2386  CAATCTCTTACAGACTCCCGACGCTGAATCTGTGAGCTCTGAGAGGAAACAAAGTACA 2445
Db      2746  TCTGTGCTTATGTGCAACCCCGAGCTGAAGAGCTTAACCTGACAGAAACCTTGAT 2805
Qy      2446  GACCAAGAGTATGCTCTCAAGATGCTTGAAGATCTCCAGCTGCGCTGACAGAAAG 2505
Db      2806  GACTTGGCGTGGCACTGCTGTGTGAGGGGCTCAGGATCTGTGCTTCAAACTATACCC 2865
Qy      2506  CTGATTAAGCA 2516
Db      2866  CTGGGGCTGGA 2876
```

```
RESULT 2
US-09-388-221B-9
; Sequence 9, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4365)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-388-221B-9
```

```
Query Match 5.5%; Score 221.4; DB 4; Length 4556;
Best Local Similarity 47.9%; Pred. No. 9.3e-55;
Matches 953; Conservative 0; Mismatches 966; Indels 72; Gaps 8;

Qy      532  TTGATTCAGACCGGTGGGCTTCGCGCAAGGTGTTCTGCAAGGAAATCAAGCA 591
Db      952  TTTGGCCCAAGGCTGGAATACCAAGAACTTCGCAATGACTGCAAGGGGCTGTGGA 1011
Qy      592  ATTGGAAATCGGCTTACGACAGAAAGATGCTGTGTGCTGTGGCGCAAGGTGACTATAC 651
```

Dh 1012 ATTGGGAAGTCAACATGCGCCAGAGGAGTGAAGAAAGCTGGGGAGAGGCCAGCTGTAT 1071  
Qy 652 CAGGGAATGTTCTCTAGCTCTTCTTCCCTCCCGTTAGAGATGACAGGGAAGAAGAG 711  
Db 1072 GGGGACCCCTTCACAGATGTCTTCTTCACTTCACTGACAGTCAAGAGCTGCCCCAGAGG 1131  
Qy 712 AGCATGTCAACAGATTCTCCAGGGAGTGGCCAGACTCCAGGGCTCCGGTGAAGAG 771  
Db 1132 GTGAGTCTGCTGAGACTCATCGGAAAAAGTGGAGACAGCACTCCGGCTCCATTGACAG 1191  
Qy 772 ATCATGTCCCAACAGAAAGCTGTGTTCATCATTTAGAGGTTTCAGTGAAGTCCGCTCT 831  
Db 1192 ATCTGTATAGGCGCAGAGGGCTGCTTTCATCTGATGTGTATGATGAGCCAGATAG 1251  
Qy 832 GTTCCCAACATGACA---CAAGCTTCGAAAGCTGGGCTGAGAGAGAGCTCCGCTTC 888  
Db 1252 GTCTTCAGAGAGCCAGATTCTGAGCTGTGTGACCTGAGACCAACAGCCGCGGAT 1311  
Qy 889 ACCCTCATAGCAGCTGTCTGAGAAAGGCTGCTCCCTGAGTCCCTTCGATGCTCAC 948  
Db 1312 GCATGCTGGGAGATTGTGCTGGGAAAACTATCTCCGAGGCACTCTTCTGATCAG 1371  
Qy 949 GTCAAGAGCTGGGAGCAGAGAAAGCTCAAGTCAAGTGTGTCTCCCGTTACCTGTTA 1008  
Db 1372 GCTCGGACCAAGCTCTGAGAACTCAATCTTCTTGGAGCAGGCAAGTTGGTAGAG 1431  
Qy 1009 GTTAGAGAAATCTCCGGGGAACAAGATCCACTGTCTCTTGAAGCGGGATTGGTAG 1068  
Db 1432 GTCTGGGGTTCTTGAAGTCCAGCAGAGAAAGATTTTCTACAGATTTTCAACATGTA 1491  
Qy 1069 CATCAGAGACAAGAGGTTGCTGATGAACCAACCGTGAAGCTGTGACAGTGC 1128  
Db 1492 AGGCAAGCAATTAGAGCTTTAGTTGGTCAATCAACAAAGACCTCTGGGCTGTGT 1551  
Qy 1129 CAGGTCCCGCGCTGTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188  
Db 1552 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1611  
Qy 1189 GAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248  
Db 1612 GAAAAAATCAACATGCTTCAAGCAAGCAACCAACCTCTGTCTTCAATTAACCTGCTG 1671  
Qy 1249 CAGCTCAACCCCTGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1308  
Db 1672 GCTTCAAGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704  
Qy 1309 AAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1368  
Db 1705 AGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764  
Qy 1369 GACGACCTCATGCTTCAAGAACTCGGGAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCT 1428  
Db 1765 GATGACCTCAAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824  
Qy 1429 ATCTCTTCCCAAGCAGCAGCTGAGAGATCACTTCACTTCTTCACTGCTGCTGCTGCT 1488  
Db 1825 A---TTCTTCAAGAGCAACCCATCTCTGAGCTTCACTTCACTTCACTTCACTTCACT 1881  
Qy 1489 GACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1548  
Db 1882 GAGTCTTCTGAGCAATGTCTTCTTCTTCTGAGAGATGAGAAAGGGAGAGTAAATCT 1941  
Qy 1549 CCTCTGTAAGTTGAGAGCAAAAGAGTCAATGAGCTTAAACAGGAGCTTCCATATC 1608  
Db 1942 AATTGATCATATGATTTGAAAAAGCTGAGAGCATATGAAATCATGAGCTGCTGCT--- 1998  
Qy 1609 CACTGCTTTGGAGTAAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1666  
Db 1999 ---GGGGATCAACCAAGTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2055  
Qy 1669 CTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728  
Db 2056 ATGGAAGAACTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2115

Qy 1729 GTCTCTGTTGGAGTCAAGAGCTTAATGCAACACCCAGAGAGACCCCTGAGGCTTC 1788  
Db 2116 TCCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2157  
Qy 1789 CACTGCTTCTTCTGAGTCAAGAGCAAAAGATTGTTGCTGCTGCTGCTGCTGCTGCTGCT 1848  
Db 2158 CACTGCTTCTGAGAGTCTGAGAGCAAAAGTCTTCTGAGAGAGTATGAGGCTTCTGAG 2217  
Qy 1849 GAAGTGGCTTCCGATTAAACAGAACTGAGCTTGAATGATCTTCTGCTGCTGCTGCT 1908  
Db 2218 GAAATGGGCAATGTGTGTAAGAAACAGACATGAGAGCTTATGCTGATCTTCTGCA 2277  
Qy 1909 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968  
Db 2278 TTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2330  
Qy 1969 TCCGCTGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2028  
Db 2331 AAGATGAGCCCAACAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2385  
Qy 2029 CAGTGGAGAAATTTCTGCTCAATGCTTGGCAACCACTGCTGCTGCTGCTGCTGCT 2088  
Db 2386 TATTGGCAGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445  
Qy 2089 GGCAGACATCTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2148  
Db 2446 AGTGAAGCTGCTGAGCACTGCAAGTGAAGAGTCTTGTGAAGACCTTGAAGAGCT 2505  
Qy 2149 ACTGCAAGATTAAGAGCCCTGATGTT---TAGAAATGACAGATTAACCTGCTGCTG 2205  
Db 2506 GCTGCTGCTGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2565  
Qy 2206 CACTTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2265  
Db 2566 GACTTGGCTTGGGCTGAGAGCCCAACAGACCTGAGCCAGCTGAGCTGAGCTGAG 2625  
Qy 2266 CACTGAGAGAGAGATGTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2325  
Db 2626 GTGCTCAAGATGCTGAGGAGCAACCTTGGCAAGAGCTGAGAGCAAGCTGAGAG 2685  
Qy 2326 TTGAGCTTGGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2385  
Db 2686 CTACAGGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2745  
Qy 2386 CAAATCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445  
Db 2746 TCTGCTTAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2805  
Qy 2446 GACCAAGGAGTATGCTTCTGATGATGCTTGAAGATCTTCCAGTGGGCTGAGAG 2505  
Db 2806 GACTTGGCTGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2865  
Qy 2506 CTGATTAAGTGA 2516  
Db 2866 CTGGGCTGGA 2876

RESULT 3  
US-09-388-221B-3  
; Sequence 3, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-1J 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 4200  
; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(4197)  
US-09-388-221B-3

Query Match 5.5%; Score 220.6; DB 4; Length 4200;  
Best Local Similarity 48.0%; Pred. No. 1.5e-54;  
Matches 939; Conservative 0; Mismatches 944; Indels 72; Gaps 8;

QY 532 TTGATTCAGACCGGTGGGGCTTCCGGCTTCGACCGGTGTTCTGACCGAAAGTCAGGA 591  
DB TTTGGCCCAAGGCTGATATCCCAAGAACCTGCATATGCAATGCAAGGGGGCTGCTGGA 1011  
QY 592 ATTGGGAAATGGGCTTAGCCAGGAAGATCGTGTGTGTGGGGGCAAGGTGAGCTGTAC 651  
DB ATTGGGAAGTCAACCTGGCCAGAGGAGGTGAGAGGCTGGGGAGAGGCCAGCTGTAT 1071  
QY 652 CAGGAATGTTCTCTAAGTCTTCTCTCCCGTTAGAGAGATGCAAGCGGAAGAGAG 711  
DB GGGGACCGCTTCCAGCATGTCTTCTAATTCAGTGCAGAGAGCTGGCCCATGCCAGGTG 1131  
QY 712 AGCAGTGCACAGATTATCTCAAGGAGTGGCCAGACTCCAGGCTCGGTGACGAG 771  
DB GTGAGTCTCGCTGAGCTCATCGAAAGATGGGACAGCACTCCGGCTCCATTAGACAG 1191  
QY 772 ATCATGTCGACGAGAAAGGCTGTGTTCATCATGACGGTTTCATGACCTGGGCTCT 831  
DB ATCTGTGTAGGCAAGACGGGCTGTCTTCACTTCAGATGTGATGATGATGACCAAGATGG 1251  
QY 832 GTCTCAACAATGACA---CAAAGCTGCAAAAGCTGGGCTGAGAAAGCAAGCTCCGTT 888  
DB GTCTTGGAGAGACGAGAGTTCGAGCTGTCTGCACTGGAGCCAGCCAGCCAGCCGGAT 1311  
QY 889 ACCCTCATACGAGTCTGTGAGAAAGTCTGTCTCTGAGTCTTCTGATGTGACC 948  
DB GCATGCTGTGGGAGATGTGGGAAACATATCTCCGAGGACATCTTCTGATACG 1371  
QY 949 GTCAAGACGTGGGACAGAGAGCTCAAGTCAAGGCTGCTGCCCGTTAATCTGTGA 1008  
DB GCTGGACACAGCTCTGCAAACTCATCTCTTGTGGAGCAGGACGTTGGGTGAG 1431  
QY 1009 GTTAGAGAAATCTCCGGGGAACAAAGATCCACTTGTCTTGAAGCGGGATGTTGAG 1068  
DB GTCCGGGGTCTCTGAGTCCAGAGAGAAATATTTCTACAGATATTTACAGATGA 1491  
QY 1069 CATCAAGAACACAAAGGTTGCGTGCATCATGAACAAAGCTGAGCTGTGACCAAGTC 1128  
DB AGGCAAGCAATTAGAGCTTTAGGTGTCAATCAACAAAGAGCTTGGGCCCTGTGT 1551  
QY 1129 CAGGTCGCCGCGTGGGCTCTCATCTGCGTGGCCGAGCTGACAGAGAGTGGTGGG 1188  
DB CTGTGGCTGGGGTCTGTGGCTGGCTGCACTTGTGATGACAGATGAGGGAAG 1611  
QY 1189 GAGAGCGTGGCCCTTCAACCAACGCTCAAGGCTGCAAGCGCTTGTGTTCAT 1248  
DB GAAGAACTCACTGACTTCAAGACACCAACCTCTGTATACATTAACCTTGGCCAG 1671  
QY 1249 CAGCTCAACCTTCAAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTGTCTG 1308  
DB GCTCTCAAGCTCAAGCCATTGG-----GACCCAGCTC 1704  
QY 1309 AAGGCTTCTGCGGATGCTGTGGAGGAGGTGGAATAGAAAGTCAGTGTGATGGT 1368  
DB AGAAGCTCTGTCTCTGTGGCTGTGAGGCACTGGCAAAAAGAACCTTTTCAATCA 1764  
QY 1369 GACGACCTCATGTTCAAGACTCGGGAGTCTGAGCTCCGTGTCTGTTCATGATGAC 1428  
DB GATGACCTCAAGAGCATGGGTAGATGGGCAATCACTCCACTTGTGAAGTGGGT 1824  
QY 1429 ATCTTCTCCAGACGCACTGTGAGGATGATCACTTCTTCCACTGATGCTCCAG 1488  
DB A---TTCCTTCAAGAGCAACCCATCTCTGAGCTACAGCTTCAATTCACCTGTGTTCCAA 1881

QY 1489 GACTTGTGCGCGCTGTACTAGTGTAGAGGCTCGAAATCGAGCAGCTCTGTC 1548  
DB GAGTTCCTTTCAGCAATGTCTTATGTCTTGGAGATAGAAAGGGAGAAAGTAAATCTT 1941  
QY 1549 CCTCTGATGTTGAGAAAGCAAGAGGTCCATGAGAGCTTAAACAGGAGCTTCATATC 1608  
DB AATTGACATATGATTTGAAAAGACGCTAGAAAGCATATGAAATATACATGCGCTGTGGG 2001  
QY 1609 CACTGCTTTGATGAAAGCTTTCTGTGTGGCTCGTGAAGCCAAAGCTAAGAGGCCA 1668  
DB GCATC-----ACCAACGTTTCTATGTGGCTGTAAAGTATGAGGGAGAGAGAG 2055  
QY 1669 CTGAGGTCCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAGAGCAAGCTTCTGACCTG 1728  
DB ATGAGAAACATCTTCACTGTCCGGCTCTTCAGGGAGAGAACTGATGTGATGCCG 2115  
QY 1729 GTCTCTGTGGGTGACAGGCTTATGCAACCAACCCAGAGAGACCTTGAAGCCTTC 1788  
DB TCCTGAGCTGTGTGACAGC-----ACACTCTGTGAGTCCCTC 2157  
QY 1789 CACTGTCTTTTCGAGACTCAAGACAAAGATTTGTGTGCTTGGATTAACAGCTTCCA 1848  
DB CACTGCTTTGACGAGCTCGGAACAAACGTTCTGACAAAGTATGGCCCATTTGAA 2217  
QY 1849 GAAGTGTGCTTCGATTAAACCAAGACCTGGAATTTGATATCTTCTTCTGCTCCAG 1908  
DB GAATGGGCAATGTGTATGAAACAGACATGAGAGCTTATGTGTGCACTTCTGATTA 2277  
QY 1909 CACTGTCTGATTTGGGAAAAATTGGGTGATGTCAAGGATCTTCCAGAGATGAG 1968  
DB TTCAGCGGCAACGGAAGAGCTTCACTGATTTGAGGGCAGG-----CAGCACATATC 2310  
QY 1969 TCCTGTAGGCAATCTGTGTGTCTTATGATGTGGGATTAAGACCTCATTTAGAGAG 2028  
DB AACATGAGGCCCAACATGTGATGTCTGTCAAGTGG3---TCCAGTCAAGATGCC 2385  
QY 2029 CAGTGGAGAAATTTCTCTCATGCTTGGACACCAACCACTGCGGAGCTGAGCTG 2088  
DB TATTGGAGATTTCTTCTCTGCTCTCAAGGTACCAAGAACTGAAAGAGCTGAGCTA 2445  
QY 2089 GGCAGACATCTGACAGAGCGGGCAGTAAAGACCTGTGTGCCAAGCTGAGGATCCC 2148  
DB AGTGAATCTGAGGACCACTGTGACGTGAAGAGCTTTGTAAAGACCTGAGAGCCCT 2505  
QY 2149 ACCTGCAAGTACGAGACCTGTATGTT---TAGAAATGCAAGATTAACCTGTGTGAG 2205  
DB CGTGGCTCTGAGAGACCTGTGGGTGGCTGTGGCTCTCAAGCTGAGAGACTGCAAG 2565  
QY 2206 CACTCTGAGAAATGTCATATGCAACCGTAACCTTAAGATCCCTCACTTGGAGGAC 2265  
DB GACTTGTGCTTGGGCTGAGAGCCCAACCAAGACCTGACCGAGCTGAGACTGTCAAT 2625  
QY 2266 CACTGAGAGAGAGATGTAGAGATGCGGTGAGACCTTAAACACCCAAATGTTTG 2325  
DB GTGCTACAGAGATGTGAGGCCAAACACTTTGCAAGAGCTGAGACCGGAGCTGCAAG 2685  
QY 2326 TTGAGATCTTTGAGGCTGAGATGTGTGATGATGACCATGCTGTATCTGGAATCTCC 2385  
DB CTACAGGAGCTGAGGCTGTGAGCTGTGAGCTCAAGTCAATGTCTCTCAGAGACTGACC 2745  
QY 2386 CAATCTTACGACCTCCCGAGCTGAAATCTTGAAGCTGAGCAAGAAACAGTGA 2445  
DB TCTGTGCTTATGTGCAACCCAGCTGATGAAGAGCTTACCTGACGAGAACCACTGTGAT 2805  
QY 2446 GACCAAGAGATGATGCTTCAATGATGCTTGGAG 2480  
DB GACGTTGGCGTGCAGCTGCTGTGAGGGGCTGAG 2840

RESULT 4  
US-09-388-221B-5  
; Sequence 5, Application US/09388221B

Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-LJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; CURRENT FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 4332  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(4332)  
US-09-388-221B-5

Query Match 5.5%; Score 220.6; DB 4; Length 4332;  
Best Local Similarity 48.0%; Pred. No. 1,66-54;  
Matches 939; Conservative 0; Mismatches 944; Indels 72; Gaps 8;  
Qy 532 TTTGATTCAGACCGGTGGGCTTCCGGCTCGACAGTGTCTGTCAGGAAAGTCAGAA 591  
Db TTTGGCCAGGCTTGATACCAAGAACCTCGCATATGATATCAGGGGGCTGCTGA 1011  
Qy 592 ATTGGAAATGGCTTACCCAGAAAGATCGTGTCTGTGGGCGCAAGTGTGACTTAC 651  
Db ATTGGGAAGTCACACCTGGCCAGGAGGTGAAGAGCTGGGGGAGAGGCGAGCTGAT 1071  
Qy 652 CAGGGAATGTTCTCTACGTTCTTCTCCGCCGTTAGAGAGTACAGCGGAAAGAG 711  
Db GGGAGCCGCTTCAGCATGTCTTCTTACCTTCACTGACAGAGAGTGGCCAGTCCAGAGT 1131  
Qy 712 AGCATGTTCAGAGATTCATCTCAGGAGTGGCCAGACTCCAGAGCTCCGGTACGAG 771  
Db GTGAGTCTGGTGAAGCTCATGGAAAGATGGGACAGCACTCGGCTCCATTAGACAG 1191  
Qy 772 ATCATGTCCGACCAAGAAAGCTGTGTTCATCATTTGACGGTTTGAATGACTGGCT 831  
Db ATCTGTCTAGGCGCAGAGCGGCTGCTTCATCTCGATGTGTAGTATGAGCAGAGATGG 1251  
Qy 832 GTTCCCAACATGACA---CAAGCTCTGCAAGAGCTGGGCTGAGAAACAGCTCCGTT 888  
Db GTTCTGAGAGCCAGATTCGAGCTGTGTGCACTGAGACCAAGCCAGCGGGAGAT 1311  
Qy 889 ACCCTCATAGCAGTCTGCTGAGAAAGTCTGCTCCGAGTCTTCCTGATGCTCAC 948  
Db GCATGCTGGGAGTGTCTGGGAAACATATATCTCCGAGGCACTCTCTGATCACG 1371  
Qy 949 GTCAAGACGTGGGACAGAAAGCTCAAGTCAGAGTGTGTCTCCCGTTACTGTTA 1008  
Db GCTGGAACACAGCTCTGCAAACTCATTCCTTTGGAGCAGGACAGTTGGGTAG 1431  
Qy 1009 GTTAGAGAAATCTCCGGGAAACAAAGATTCACCTGTCTCTTGGAGCGGGATGAG 1068  
Db GTTCTGGGGTCTCTGAGTTCAGCAGGAAAGATTTCTTACAGATTTTTCACATGAA 1491  
Qy 1069 CATCAGAGACCAAGGGTGTGTCATGAAACCAAGTGTGAGCTGTGAGCTGAGCAGTGC 1128  
Db AGGCAAGCAATTAGAGCTTTTAGTGTGATTAATAAACAAAGACTCTGGGCGCTGTGT 1551  
Qy 1129 CAGGTGCGCGGTGGCTCTCTCATCTGCTGAGCGCTGACAGTGTGAGCG 1188  
Db CTGTGGCTGGGTGTCTGCTGGCTGAGCTGCTGATGACAGATGAGGAG 1611  
Qy 1189 GAGAGCGTGGCGCTTCAACCAACGCTCAAGGCTGACAGCGGCTTTGTGTTCAT 1248  
Db GAAAAACTCACATGACTTCAGAGCAACCAACCTCTGTCTACATTAATCTTGGCCAG 1671  
Qy 1249 CAGCTACCCCTCGAGGCGGTGGCTGTCTCATATGAGAAAGATGTCTG 1308  
Db GCTCTCAAGCTCAGGCAATGG-----GACCCAGCTC 1704

Qy 1309 AAGGCTTCTGCGGTGAGGAGGTGGAATAGAAAGTCAGTGTGATG 1368  
Db AGAGACTCTGTCTCTGGCTGTGAGGAGCATCTGGCAAAAGACCTTTTCACTCA 1764  
Qy 1369 GACGACTCATGATTCAGAGACTCGGGAGTCTGAGCTCGTGTCTTTCATGAAAC 1428  
Db GATGACCTCAGAGAGCATGGGTATGATGGGCAATCATCTCACCTCTTGAAGATGGT 1824  
Qy 1429 ATCTTCTCCAGACAGCACTGTGAGGATCTACACTTCTTCACTCAGTCTCAG 1488  
Db A---TTCTTCAAGAGCAACCCCATCTCTGAGCTTACACTTCACTCAGTCTTTCAA 1881  
Qy 1489 GACTTCTGCGCTGTGATGATGATGAGGCGCTGAGAAATGACCAAGCTCTG 1548  
Db GAGTCTTTCAGCAATGTCTTATGTCTTGAAGATGAGAGGAGAGATTAATCTT 1941  
Qy 1549 CCTGTACGTTGAGAAACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCATATC 1608  
Db AATTGATCATATGATTTGAAAGAGCTGAGAGCATATGAAATATCATGCTGTTGG 2001  
Qy 1609 CACTGCTTTGATGAAAGCTTTCTTTTGGCTGTGAGCGAAGACGTAAAGAGCCA 1668  
Db GCATC-----AACCAACGTTTCTTATTTGGCTGTGATGATGAGGGGAGAGAG 2055  
Qy 1669 CTGAGAGTCTGCTGAGGCTGCTCCGTTCCCTGGGGGGTGAAGCAGAACTTCTGACTGG 1728  
Db ATGAGAAACATTTTCACTGCGGTGTCTGAGGGAGAAACCTGATCAGTGTCCG 2115  
Qy 1729 GTCTCTGTGTTGGGTGACAGCTTAAATGCAACACCCAGAGAGACCTCGAGCCCTTC 1788  
Db TCCCTGAGCTGTGCTGAGAC-----ACACTCTGTGAGTCCCTC 2157  
Qy 1789 CACTGTCTTTTGAAGACTCAAGAGAGTTTGTGCTTGGCATTAAGACTTCCAA 1848  
Db CACTGCTTTGAGAGACTCGAGACAAAGCTTCTGACACAAAGATGAGCCATTGCA 2217  
Qy 1849 GAACTGTGCTTCCGATTAACAGAACTGAGATTTGATGATCTTCTTGTCTCCAG 1908  
Db GAAATGGGCAATGTGTGAGAAACAGACATGAGAGCTTTAGTGTGACATTTTCA 2277  
Qy 1909 CACTGTGCTTTTGGGAAATTCGAGGTGATGCAAGAGGATCTTCCAGAGATGAG 1968  
Db TTCAAGCCCGCAGTGAAGAGCTTCACTGATTAAGGAGG-----CAGCAGATC 2330  
Qy 1969 TCCGCTGAGGATGCTGTGTCTCTTATGATGTCGGGATTAAGACCTTATGAGAG 2028  
Db AAACATGAGGCCCAACATGATGATGCTGTGAGGTGG-----TCCAGTCAAGATGCC 2385  
Qy 2029 CAGTGGAGAAATTTGTCTCATGCTGTGACCAACCACTGCGGAGCTGAGCTG 2088  
Db TATTGGGAGATTTCTTCTCGTCTCAAGGTCAACAGAAACCTGAGAGAGCTGAGCTTA 2445  
Qy 2089 GCGAGCAGATCTGACAGAGCGGGCATGAAAGCCGTGTGCAAGCTGAGGATGCC 2148  
Db AGTGAAGCTGCTGAGCACTGTGAGTGAAGGTCTTTTGAAGACCTTGAAGCGCTT 2505  
Qy 2149 ACTGCAAGATACAGACCTGATGTT---TAGAATGACACAGATTAACCTCTGTGTGAG 2205  
Db CGCTGCTCTGAGAGACCTGCGGTGTGCTGTGCTGCTCAAGCTGAGAGAGCTGCAAG 2565  
Qy 2206 CACTGTGAGAAATGCTATGAGCAACCTTAACCTTAAGATCTTCACTTGGAGGAC 2265  
Db GACCTGTGCTTTGGGCTGAGAGCCAAACAGACCTTGAACCTGAGCTTCAAT 2625  
Qy 2266 CACTGAGAGAGAGATGTAGAGTGGCGTGTGAAGCTTAAACCCCAAAATGTTG 2325  
Db GTGCTCAGAGATGCTGAGCCAAACCTTTGCAAGACTGAGACAGCCAGCTGCAAG 2685  
Qy 2326 TTGAGTCTTTGAGGCTGATTTGCTGTGATGACCAATGCTGTAACTGAAAGATCTCC 2385  
Db CTACAGGAGACTGAGCTGTGAGCTGTGAGCTTCAAGTGTGCTGTGCAAGACTGAGCC 2745

Qy	Db	Qy	Db
2386	2746	2446	2806
AAATTCCTTACGACCTCCCGAGCTGAAATCTCTGAGCTTCGACGAGAAACAAGTGACA	TCTGTCTTAACTGCTCAGCCGACCTCGAAGAGACTGAGACTTCGACGAGAAACAACCTGGAT	GATCCAGGAGATATGCTCTTCTCAGTGATGCTTGAAG	GACGTGGGTGTCGACTGCTCTGTGTAGGGGCTTAC
2445	2805	2480	2840
AAATTCCTTACGACCTCCCGAGCTGAAATCTCTGAGCTTCGAGCTTCGACGAGAAACAAGTGACA	TCTGTCTTAACTGCTCAGCCGACCTCGAAGAGACTGAGACTTCGACGAGAAACAACCTGGAT	GATCCAGGAGATATGCTCTTCTCAGTGATGCTTGAAG	GACGTGGGTGTCGACTGCTCTGTGTAGGGGCTTAC

## RESULT 5

```

US-09-388-221B-11
; Sequence 11, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 4466
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4272)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-388-221B-11

```

Query Match	5.5%	Score 220.6	DB 4	Length 4466
Best Local Similarity	48.0%	Pred. No. 1.6e-54		
Matches 939, Conservative	0	Mismatches 944	Indels 72	Gaps 8

Qy	532	TTTGATTTGACGCGGGTGGGGCTTTCCGGCGCTGCAACGGATGTTCTGCAAGAAAGTCAGAA	591
Db	952	TTTGGCCAGGCGCTGGATATCCAGAACCTGGCATATGTCATATCTCAAGGGGGCTCTGGA	1011
Qy	592	ATTGGGAATCGGCTCTAGCCAGAAAGATCGTGTGTCTGGGCGCAAGTGGACTTAC	651
Db	1012	ATTGGGAAGTCAACACTGGGCCAGGCGAGGTGAAGAGCCTGGGGGAGAGGCCAGCTGTAT	1071
Qy	652	CAGGAATGTTCTCTAGGTCTTCTCCGCCCGTTAGAGAGATGCAACGGAAAGAG	711
Db	1072	GGGAGCGCTTCCAGCATGTCTTCTTACTTCACTGACCTGCAAGACGTGGCCAGTCCAAAGTG	1131
Qy	712	AGCAGTGTACAGAGTATCTTCCAGGAGTGGCCAGATCCCAAGCTCCGGTACGAG	771
Db	1132	GTGAGTCTCGCTGAGCTCATCGGAATAATGGGACAGCACTCCGGCTCCCATTAGACAG	1191
Qy	772	ATTCATGTCGCCACCGAAGAGCGTGTGTTCAATCTTAAGGTTTGAGATGACCTGGGCTCT	831
Db	1192	ATCTGTTTAGGCGCAGAGCGGTGCTCTTTCATCTCTCAGATGGTATGATGACCAAGATGG	1251
Qy	832	GTCTCTCAACAATGAC---CAAGCTCTGCAAAAGACTGGGCTGAGAACAGAGCTCCGTTTC	888
Db	1252	GTCTTTGACGAGGCCGAGTTCTGAGCTCTGTCTGCACTGGAAGCCAGCCACAAGCCGGCGAT	1311
Qy	889	AACCTCTATACGCAAGTCTGCTGAGGAAAGTCTCTGCTCCTGAGTCTTCTCTGATCGTCACC	948
Db	1312	GCACGTCTGGGAGAGTTCTGTGGGGAAACCTAATCTTCCGAGGACATCTTCTCTATATACG	1371
Qy	949	GTCAAGAGACGTGGGACAGAGAGCTCAAGTCAAGTCAAGTCTGCTGCCCGTTACCTGTATA	1008
Db	1372	GCTGGGACCAAGAGCTGTGAGAACTCATATCTCTTTTGGAGCAAGCAAGTGGGTAGAG	1431
Qy	1009	GTTAGAGGAATCTCCGGGGAAACAAGATCCACTTTCCTCTTGGAGCGGGGATTTGGTAG	1068
Db	1432	GTCTCTGGGATTTCTGAGGTCCAGCAGGAAAGAAATTTCTTACAGATATTTTCAACATATA	1491

OY	1069	CATAGAAAGACAAAGGGTTCGGTGGATATATAGAAACAACGTAGAGCTGTGCACATGCG	1128
Db	1492	AGGCAAGCAATTAGAGCTTTAAGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGT	1551
OY	1129	CAGGTGCCCGCGGGGCTCTCTCATCTGTGGGGCCCTGAGAGCTGCAGAGCGTGGTGGG	1188
Db	1552	CTTGTCCTTGGGTGTCTGGCTGGCCGTGCATCTTGCTGTATGCAAGATGAAGGGGAAG	1611
OY	1189	GAGAGCGTCCGCCCTTCAACCAAGGCTCAAGGCGTGCACGGCGCTTTGTGTTTCAAT	1248
Db	1612	GAAGAACTCAACATGACTTCCAGAACCAACAACCTCTGTCTCAATTACCTTGGCCAG	1671
OY	1249	CAGGTCAACCCCTGAGGCGTGTGTCGGGCGGTCTCTAATCTGAGAGAAAGTTGTCTG	1308
Db	1672	GCTCTCCAGTGCACCAATTGG-----GACCCAGCTC	1704
OY	1309	AAGCGTTTCTGCCATATGCTGTGAGAGAGATGAGAAATAGAACTCAGTGTTCATGTGT	1368
Db	1705	AGAACTCTGTCTCTTGGCTGTGAGAGGATCTGGCAAAAAAGACCCTTTTCAGTCCA	1764
OY	1369	GACGACCTCATGTGTTCAAGGACTGGGGAAGTCTGAGCTCCGTGCTGTGTTTCACATGAAC	1428
Db	1765	GATGACCTCAGGAAGCAGTGGGTATGATGAGGGCCATCATCTCCACCTTCTTGAAGATGGGT	1824
OY	1429	ATCCTCTTCCCAAGCAGCCACTGTGAGAGATCTACACCTTTCTTCCACTTCAGTCTCCAG	1488
Db	1825	A---TTCTTCAAGAGCAACCCCATCCCTCTGAGCTACAGCTTCATTCACCTGTGTTTCAA	1881
OY	1489	GACTTCTGTGCGGCTTGTACTAGTGTATGAGGGCGCTGGAATCGAGCCAGCTCTCTGCG	1548
Db	1882	GAGTTCTTTGCAGAAATGTCTTAATGTCTTGGAGGATGAGAAAGGGAAGAGTAAACATCTT	1941
OY	1549	CCTCTGTACGTTGAGAAAGCAAGAGGTCCATGAGGCTTAAACAGGAGGCTTCCATATC	1608
Db	1942	AATTGCATCATATGATTTGAAAAAGACCTAGAAAGCATATGGAATATCATGGCGTGTGGG	2001
OY	1609	CACCTCGCTTGGATGAAGCGTTTCTGTGTGGCCTCGTGAACCAAGCGTAAAGAGGCCA	1668
Db	2002	GCATC-----AACCAACAGTTTCTTAATGGGCGTGTAAAGTATGAGGGGGAAGAGAG	2055
OY	1669	CTGAGAGGCTCGCGGGCTGTCCCGTCCCTCGGGGGTGAAGCAGAGCTTCTGCATCGG	1728
Db	2056	ATGAGAAACATCTTTCACCTGCGGCTGTCTCAGGGGAGAACTGTATGCAAGTGGGTCCG	2115
OY	1729	GTCCTCTGTGTGGGTCAAGACCTTAATGSCAACACCCAGAGAGCAACCTTGACGCTTC	1788
Db	2116	TCCTGCAAGCGTGTGTCGAGGCC-----ACACTCTGTGAGGTCCCTC	2157
OY	1789	CACGTGCTTTTTCGACTCAAGACAAAGAGTTGTTCGCTTGGCATTTAAACAGCTTCCA	1848
Db	2158	CACGTCTGTATCGAGACTCGGAACMAAACGTTCTGTGACAAAGATAGGCCCATTTTGAA	2217
OY	1849	GAACTGTGCGTTCGATTAACAGAACCTGSACTTGATAGCATTTCTCTGTGCTCCAG	1908
Db	2218	GAATATGGGCATGTGTGTAAGAAACAGACATGAGCTCTTAAGTGTCACCTTCTGTGATTTAA	2277
OY	1909	CACGTCCGTAATTTGCGGAAAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAGATGAG	1968
Db	2278	TTTACGCCGCCACGTGAAGAAAGCTTCAGCTGATTTAGGGGACG-----CAGACACATC	2330
OY	1969	TCCGCTGAGGCAATGTCTGTGTGTCCTCTATATGATCGGAGATTAAGACCTTCATTTAGAGAG	2028
Db	2331	AACATGAGGCCCAACATGTAGATGTCGTTCAGGTGG-----TCCAGTTCACAGATGCC	2385
OY	2029	CAGTGGGAAGATTTCGTCTCATATCTTTGGACCAACCAACCTGTGGGCAAGCTGGACCTG	2088
Db	2386	TATTGGAGATTCTCTCTCCGTCTCAAGGTCAACGAAGACCTGAAGAGGCTGGACCTTA	2445
OY	2089	GGCAGCAGCAATCTGCAGAGAGGGGCATGAAGACCTGTGTGTCCAAGCTGAAGCATCCC	2148
Db	2446	AGTGGAAATCTGCTGAGGACACTCTGCAGTGAAGAGTCTTTGTATGAAGACCTGTGAACGGCCT	2505
OY	2149	ACCTGCAAGATACAGACCTGTATGTT---TAGAAATGCAACAGATTACCCCTGTGTGACG	2205





Accession	Sequence	Position
Db	1062 GGAGCTGTGGCCAGGCGCTTGAGCCGCGGACCAACACTGGGGGGCTCTGTCTTTGGGGA	1121
Qy	3033 ATGTGGACTGACTTCTGATTTCTGTGAGGCACTCTCTTGGGCGCTTTCCTGCAACCGCA	3092
Db	1122 CTGTGAGGTGACCAACAGCGGCTGAGCAGCGCTGCGCTCTCTGCTGGCCAAACCGAG	1181
Qy	3093 TCTGACCAGTCTAAACCTGTGTGCAGAAATTACTTCACTCCCAAAGAAATGATGAACCTGTG	3152
Db	1182 CCTCGAGAGCTGGGACCTGAGCAACATCTGTGTGGCGACCGCGGCGCTCTCGACAGCTGCT	1241
Qy	3153 TTTCGACCTTTGCTCTGTCTCCACAGCTTAACTTACAGATTAATTGGGCTGTGTGAATGACAGTA	3212
Db	1242 GGGAGACCTGGAGCAGCGCGGGCTGGCGCTTGGAGCAGCTGGTCTGTATCGACACACTACTG	1301
Qy	3213 CCTGTGTCAATTAAGGAAGCTGTGTGAGGAAGTGCAGCTACTCAAGCCCCG	3263
Db	1302 GACGAGGAGGTGAGGACCGCTCTGACAGCCCTTGAAGGGAGCAAGCCCGG	1352

Db 1182 CTTGCGAGAGCTGAGACCTTGAGCAACAATGTGTGGCGACCCGGGCTCTGAGCTGCT 1241  
QY 3153 TTTCGCGCTTTCCTGTCCTCCACGCTCTTAATTGAGCTGTGGAATGAGCA 3212  
Db 1242 GGGGAGCCTTGAGACACCGGGCTGCGCCCTTGAGAGCTGTCTTACGACCTACTG 1301  
QY 3213 CCTGTGCAAAATAGAGAGCTGTGAGAGAGTGCAGTACTAGACCCCG 3263  
Db 1302 GACGAGAGAGTGTGAGAGACCTGCTGACAGCCCTTGAGGGGAGAGACCCCG 1352

RESULT 8  
US-08-910-731-5  
Sequence 5, Application US/08910731  
Patent No. 5932440  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
APPLICANT: SHANDILYA, HARINI  
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,731  
FILING DATE: (Herewith)  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: 08/795,395  
FILING DATE: 04-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/794,546  
FILING DATE: 03-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0942.3440003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2600  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1386 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
US-08-910-731-5

Query Match 4.5%; Score 182.4; DB 2; Length 1386;  
Best Local Similarity 47.1%; Pred. No. 1.9e-43;  
Matches 595; Conservative 0; Mismatches 666; Indels 3; Gaps 1;

QY 2000 GGATGGGGATTAAGACCTCATTTGAGGAGAGTGGGAGAGATTTCGTCTCATGCTTGGCA 2059  
Db 101 GGCTGAGCAGCTGTGGCTCTCAAGAAAGCAGGTGCAAGACATGAGCTTGCATCTTCGAG 160  
QY 2060 CCCACCCACACCTGCGGAGCTGGAACCTGGGAGAGAGCATCTGACAGAGCGGGCCATGA 2119  
Db 161 TCACCTTGACATCGGAGAGAGCTCAACCTTGCGAGCAAGAGCTGGGCGATGTGCGGCTGC 220

QY 2120 AGACCCCTGTGTGCAAGCTGAGGACATCCACCTGACAGATACAGACCCGATGTTTAA 2179  
Db 221 ATTGGGTGTCCAGAGGCTGTCAGAGCCCTCTCTGAAAGATCCAGAACTGAGCTTCAGA 280  
QY 2180 ATGCAAGATTACCC---CTGTGTGACAGACCTTGTGAGAAATCGTATGGCCAACTGA 2236  
Db 281 ACTGTGCTGACGGGGGGCGGCTGCGGGGTCCTGTCCAGACACTAGACACCTTGCACCA 340  
QY 2237 ACTTAAAGATCCCTCAACTTTGGAGGACCCACCTGAAGAGAGATGTAAGATGGCGT 2296  
Db 341 CCTTCAGAGAGCTGACCTCAGACCAACCTCTTGGGGGATGCGGGCCCTGACGCTCT 400  
QY 2297 GTGAAGCTTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGATTCGTGTGAT 2356  
Db 401 GCGAAGAGCTCTTGAGACCCCACTGCCCTGCTGAAAGCTGACGCTTGAATTTGACGCC 460  
QY 2357 TGACCCATGCTCTGTTTACCTGAAGATCTCCAAATCTTACGACCTTCCCGACCTGAAT 2416  
Db 461 TCTGGGCTGCGAGCTGCGAGGCCCTGGGCTCGTGTCTCAGGGCCAAAGCCGACCTCAAG 520  
QY 2417 CTCTGAGCTTGAGAGAAACAAAGTGCAGACACAGAGATTAATGCTCTCACTGATGCTT 2476  
Db 521 AGCTCAGGTTTGAACAACGACATCAATGAGGCTGGGCTGCTGTGTGCTGCTCCAGGGCC 580  
QY 2477 TGAGAGCTTCCAGTGGCGCCCTGCGAAGCTGATCTGAGAGGACCTGTGCAATCAGACCA 2536  
Db 581 TGAAGGACTTCCCTGCTGCGAGCTGAGAGGCGCTCAAGTGAAGCTGCGGTGGAATCAG 640  
QY 2537 CGGGTTGCGCAGAGTGTGCGCTTCAAGCCCTGTCAGAGAACCGGAGTTTGAACAACCTGTGCG 2596  
Db 641 ACAAGTCCGGGACCTGTGCGGACATTTGTGCTCTCAAGGCTCGGTGCGGAGAGCTGCGCC 700  
QY 2597 TATCCAAACAACCTGCGGAGAAAGAGGTGTAATCTATCTGTGTGATTCATGAGAGCTTC 2656  
Db 701 TGGGAGAGCAACAGCTGGGTGATGTGGGACATGCGAGCTGTGCGCAGGGCTGTCCAACC 760  
QY 2657 CCCACTGTATGTGAGAGGCTGATGTGATCACTGCGACCTGTGACACAGCTGTGCTGTG 2716  
Db 761 CAGCTCAGAGGTCAAGACCTGTGATCTGGAATGTGGATCACTGCAAGGAGGCTGTGCG 820  
QY 2717 GTTTTCTTGACATTTGCGCTTAAGGGTAACTCATATGAGCTGACGCACTGAGACCTTACATGA 2776  
Db 821 GGGATCTGTGCGGTGTCTCTCAAGGCCAAGAGAGCTTAAGAGCTTCAAGCTTGGCGGCA 880  
QY 2777 ACCCTGTGAGAGACAATGCGCTGAAGCTTCTGTGAGAGTCAATGAAGAACCATTTGTC 2836  
Db 881 ACGAGCTGAGGAGATGAGGGGTGCGGACATGCTGTGTGAACCTGTGTGAACCTGCTGCC 940  
QY 2837 ATCTCAAGACCTGAGATTTGTGAAGTCAATCTTACCGCGGCTGTGTGAGATCTGT 2896  
Db 941 AGCTGAGTGTGCTGTGGGTGGAAGTCTTCAAGCTTCAACGCGCTGTGCGCCCACTTCA 1000  
QY 2897 CCTGTGATCTCGAGAGAGACAGACACTGAAGAGCTGAGATCTCAAGGCAATGCGCTGG 2956  
Db 1001 GCTCAGTGTGCGCCAGAAAGGTTTCTCTGAGACTACAGATTAAGCAACAGGCTGG 1060  
QY 2957 GTGACGTTGGGGTTGCTGCTGCTGTGCGAGGAGACTGAACAAAGAAAGAGTGTTCAGCA 3016  
Db 1061 AGGATGGGGGGGTGCGGAGGCTGTGCGAGGGCTGGGGCAAGCTGTGCTGTGCGGGG 1120  
QY 3017 GACTGGGTTGAAGGATGTGGACTGACTTCTGATTTGCTGTGAGGCACTTCTTGGGCC 3076  
Db 1121 TGCTTGTGTGGCCGACCTGCAATGTGATGACAGAGCTGAGAGGCTTCCCGCAACCC 1180  
QY 3077 TTTCTGCAACCGGATCTGACAGCTTAAACCTGTGAGAGATTAATTAAGTCCCAAG 3136  
Db 1181 TGTGTGCAACACAGCTGTGAGCTGAGCTTCAAGAAACAATGCTTGGGGAGAGCGCG 1240  
QY 3137 GAATGATGAAGCTGTGTGCGCTTGTGCTGTGCTTCCAGCTTAACTTACAGATTAATTTGGGC 3196  
Db 1241 GCATCTTGACATGTGTGAGAGCGTCCGGGAGCGGCGGTGCTCTGTGAGAGAGCTGTGCTC 1300  
QY 3197 TGTGAAATGAGCAGTACCTGTGTGAATTAAGAAAGCTGTGAGAGAGTGCAGCTACTCA 3256



## RESULT 10

US-08-910-731-3

Sequence 3, Application US/08910731

Patent No. 5932440

GENERAL INFORMATION:

APPLICANT: CHATTERJEE, DEB K.

APPLICANT: SHANDILYA, HARINI

TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSER: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., N.W., SUITE 600

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/910,731

FILING DATE: (Herewith)

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/795,395

FILING DATE: 04-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/794,546

FILING DATE: 03-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/024,057

FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: ESMOND, ROBERT W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.3440003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1374 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1368

US-08-910-731-3

Query Match

Best Local Similarity 4.4%; Score 175.8; DB 2; Length 1374;

Matches 568; Conservative 0; Mismatches 632; Indels 3; Gaps 1;

Db

2000 GGATCGGAGTAAAGCCCTGATGAGAGAGTGGGAGATTTGCTGCTGATGAGCA 2059

Db

86 GGCTGATGATCTGTGGCTCTCACTGAGTGGGTCGAGTCAAGTCAAGTCAAGTCAAG 145

Db

2060 CCCACCAACCTGCGGAGCTGAGACTGGGAGAGCATCTGACAGAGCGGCGCATGA 2119

Db

146 CCAACCTGCTGCTGAGAGTCAAGCTGACGCTACGACCAATGATCTGGGTGATGCTG 205

Db

2120 AACCTGTGTGCTGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 2179

Db

206 GTCTGTGTGCTGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 265

Db

2180 ATGACAGATTAACCC---CTGTGTGAGCAACCTGAGAGATGCTGATGAGCAACCTGA 2236

Db

266 ACTGACGCTTGAAGAGTGTGGGTCTGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCT 325

QY 2237 ACCTAAGATCCCTCAACTTGGGAGCACCCACTGAAGAGAGATGTAAGATGGCGT 2296

Db 326 CCTGCGTGAATCAATCTCAATGACAACTCTGGGAGTGAAGCTGAGCTGCTT 385

QY 2297 GTGAAGCCTTAACCAACCAAAATGTTGTTGAGTCTTTGAGCTGATGTTGCTGAT 2356

Db 386 GTGAAGAGATCCGGGAGCCCCAGTCCGCTTTGAGAACTTCAAGTTGGAATCTGTAAC 445

QY 2357 TGACCAATGCTGTTAAGCTGAAATCTCCAAATCTTACAGACTTCCCGAGCTGAAT 2416

Db 446 TCACAGTACCAAGCTGAGCCCTGAGCTGAGTCTGAGGATGAACCTGACTTTAAAG 505

QY 2417 CTCTAGCTGAGAGAAACAAAGTGAAGACAGAGAGTATGCTTCAATGATGCT 2476

Db 506 AGCTGATATGAGCAATGATCTTCATGAGGCTGTATCCACCTGTGTCAGGAGCC 565

QY 2477 TGAGAGTCTCCAGTGGCCCTGAGAGCTGATCTGAAGAGCTGTGGCATCAGAGCA 2536

Db 566 TGAAGATTTGCTGCTGCTCACTGAGTCACTCAAACTGAGAACTGTGTATCAGTCA 625

QY 2537 CGGTTGCTGAGAGTGGCTTCAAGCTGCTGAGCAACCGAGCTTGAACACTGTGCC 2596

Db 626 CCAACTGCAAGGATCTGTGATGTTGGCTTCAAGCTCACTGAGCAAGTCAAGTCA 685

QY 2597 TATCCAAACAGCTGGGAGCAAGAGTGTAAATCTCTGTGATCCATGAGGCTTC 2656

Db 686 TGGGAGCAAGAGTGTGGGCAACAGGCAATTCAGCACTGTGCTCAGAGCTGCTGCTTC 745

QY 2657 CCCAGTGTAGTGTGAGAGGCTGATGCTGAATCAGTGCACCTGAGCAACGCTGAGCTGTG 2716

Db 746 CCACTGAGAGTGTGAGAGCTGTGAGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 805

QY 2717 GTTTCTTGAACCTGCGCTTATGAGTAACTGATGCTGAGCTGAGCAACCTGAGCTT 2776

Db 806 AGAAGCTGAGCTGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 865

QY 2777 ACCCTGAGAGCAATGAGCTGAGAGCTTCTGTGAGAGTGTGAGAGAGTGTGAGAGTGTG 2836

Db 866 ATGAGCTGAGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 925

QY 2837 ATCTCAGAGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 2896

Db 926 AGCTGAGTCACTGTGAGTGAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 985

QY 2897 CCTGTGTATCTGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2956

Db 986 GCTGTGTGTGAGCAAAAGAGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1045

QY 2957 GTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3016

Db 1046 GAGACTCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1105

QY 3017 GACTCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3076

Db 1106 TGTCTTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1165

QY 3077 TTTCTGCAACCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3136

Db 1166 TGTGTGCAACCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1225

QY 3137 GAATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3196

Db 3197 TGT 3199

Db 1286 TGT 1288

## RESULT 11

US-08-795-395-3

Sequence 3, Application US/08795395

Patent No. 5965399  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
APPLICANT: SHANDILYA, HARINI  
TITLE OF INVENTION: Cloning and Expression of Rat Liver and  
TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,395  
FILING DATE: 04-PEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0942.3440002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1374 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1368  
US-08-795-395-3

Query Match 4.4%; Score 175.8; DB 2; Length 1374;  
Best Local Similarity 47.2%; Pred. No. 1.7e-41;  
Matches 568; Conservative 0; Mismatches 632; Indels 3; Gaps 1;  
QY 2000 GGATCGGGGATTAAGACCTCTCATTTGAGAGAGAGTGGGAAAGATTTCTGCTCCATGCTTGGCA 2059  
DB 86 GGGTGAAGAGCTGTGGCTCTCACTGAAGTGGGTGAAAGATCAAGGTCAAGCATCCAG 145  
QY 2060 CCCACCAACCTGCGGAGCTGAGCTGAGCTGGGAGAGCATCTGACAGAGCGGCGCATGA 2119  
DB 146 CCAACCTGCGCTGACAGAGCTCAGCTACGCAACATGAATGAGTGTGATGCTGGTGG 205  
QY 2120 AGACCTGTGTGCCAAGTGAAGCATCCCACTGCAATACAGACCTGTGATGTTAGAA 2179  
DB 206 GTCTGTCTCTCGAGGCTGAGAGATCCCACTTGTAAATCCAGAGGCTGAGCCTTCAGA 265  
QY 2180 ATGCACAGATTACCC--CTGGTGTGACAGACCTCTGAGAAATCGTCAATGGCAACGCTA 2236  
DB 266 ACTGAGAGTTAGCGAAGCTGTGGGTGTGGGTCTCTGCTGATGTCTGCTTTGTCTTA 325  
QY 2237 ACCTAAGATCCCTCACTTGGAGGACACCACTTGAAGAAAGAGATGTAAGATGGCGT 2296  
DB 326 CCTGCGTGAATCAATCATGACCAACCTCTGGGAGATGAAGGCTGAAGCTGCTCT 385  
QY 2297 GTGAAGCTTAAACACCCAAATATTTGTTGATGATCTTTGAGGCTGATGCTGTGAT 2356  
DB 386 GTGAAGAGCTCGGGACCCCAAGTCCGTCTTGAAGACCTTCAATTGAATATGTAACC 445

QY 2357 TGACCAATGCTGTATTAACGTGAAGATCTCCAAATCCTTAAGACCTCCCGAGCTGAAT 2416  
DB 446 TCACAGCTACAGCTGAGAGCCCTGGCTCTCAGTGTCAAGGTGAACCTGACTTAAAG 505  
QY 2417 CTCTGAGCTGGCAGAAACAGGTGACAGACAGGAGATTAAGCTCTCTCATGATGCT 2476  
DB 506 AGCTAGATTGAGCAACAAATGACTTCATGAGGCTGTATCCACACTGTGCGCAGGGCC 565  
QY 2477 TGAAGTCTCCAGTGGGCCCTGAGAAAGCTGATACGTGAAGAGCTGTGGATCAACAGCA 2536  
DB 566 TGAAGATTCTGCTCTCAACTGAGATCACTCAACTGGAAGAACTGTGTATACATCAG 625  
QY 2537 CGGGTTCACAGATCTGAGCTCAAGCCCTGTCAAGCAACCGAGCTTGACACCTGTGGCC 2596  
DB 626 CCAACTGCAAGATCTGTGTGATGTGTGGCTCAAGGCTCACTGCAAGAACTGGAAGT 685  
QY 2597 TATCAACACAGCTTGGGAAACGAGGTGAATTAATCACTGTGTGATCAATGAGCTTC 2656  
DB 686 TGGCAGCAACACAGCTGGGCAACAGGCAATTCAGCACTGTGCTCAGAGCTGCTTC 745  
QY 2657 CCCACTGTATCTGACAGAGGCTGATGCTGAATCAGTCCCACTGGAACAGGCTGGCTGTG 2716  
DB 746 CCACTGAGGCTGAGACTCTGTGGCTCTGGAGCTGTGATGTCACTGACAGAAAGCTGCA 805  
QY 2717 GTTTCTTTCGACTTGCCTTATGGGTAACTCATGCTGAGCACTGAGCTTATAGATGA 2776  
DB 806 AGGACTGTGCTGTGTCTCAGAGCCAGAGAGCTGAAGAACTCAGCTGCTGCTGGCA 865  
QY 2777 ACCCTGGAAGACATGCGGTGAAGCTTCTGTGCGAGTGTCAATGAGAAACATCTTGTG 2836  
DB 866 ATGAGCTGAAGATGAGGGTGCCCACTGCTGTGTGAGAGCTGTGTAGAGCTGGCTGTG 925  
QY 2837 ATCCCAAGACCTGGAATGTGGTAAAGTGTCACTCAACGCGGTGTGTGAGAGTGTGT 2896  
DB 926 AGCTGAGTCACTGTGGTAAAGCTGTAGCTTCAAGCTGCTCTTGTGCCACTTCT 985  
QY 2897 CTTGTGTATCTCGAGAGAGACACCTGGAAGCTGTGATCTCACAGCAATGCTGCTGG 2956  
DB 986 GCTGAGTGTGACCAAAAAGATTCTGTGTTGATGGAATGAGCAGCAACCGGTGG 1045  
QY 2957 GTGACGTTGGGTTTGTGCTGCTGTGCGAGGAGCTGAAGCAAAAGAACATGTTTGAACA 3016  
DB 1046 GAGACTGGGAGTGTGAGCTTGTGCAAGGCCCTGGGCTATCCGGAACAGTGTGCTGG 1105  
QY 3017 GACTCGGTTGAAGAGATGGAAGTGAAGTCTGATTTGCTGTGAGGAGCACTGCTTGGCCC 3076  
DB 1106 TGCTTTGCTGGGAGCTGTGATGTGACAGACAGTGTGAGCAGCTTGGCACTGTCT 1165  
QY 3077 TTTCTGCAACCGGCACTGTGACAGCTTAAACCTGTGAGAAATTAATCTGATCCCAAG 3136  
DB 1166 TGCTGGCAACCGGAGCTTGAAGGAACTGGAACCTCACTGATTAACAATGAGGAGACA 1225  
QY 3137 GAATGATGAAGCTGTGTGCGCTTTGCTGTGCTGCTGCAAGTCTTAACATTAAGTAATG 3196  
DB 1226 GTGTCTTAACAATGTCTGAGAGGCTTAACAAGCCAGCTGATCTTCAAGAGCTGTCC 1285  
QY 3197 TGT 3199  
DB 1286 TGT 1288  
RESULT 12  
US-08-232-463-14/C  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHRIFFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner



STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ19+Fls  
US-08-232-463-14

```

: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 16600
: LENGTH: 298336
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(298336)
: OTHER INFORMATION: n = A,T,C or G
: US-09-949-016-16600

```

Query Match	1.3%	Score 51.4;	DB 4;	Length 298336;
Best Local Similarity	58.0%;	Pred. No. 0.011;		
Matches 91; Conservative	0;	Mismatches 66;	Indels 0;	Gaps 0;

QY	3273	TTGAGCGTAGTTGGACATCTTTTGATGAAGTACCGCACAAAATAGGACTTACTTTCCG	3332
Db	202167	TTAGAGAAATGTGTGATTATTATTATTCATTCAGCATGACATACCAAGGTAGTTGGCATTTCCT	2023226

QY	3333	GCTCCCTGAAGCCGGGCAATGCCATCTGCTTGTGGGGGATGAACCCAGACAGAA	3392
Db	202227	ATTCTCTATACATTAAAGCATTAGCCATGTAACTTGTCTGTGCAATGAATCATGAAGAAGAA	202286

Qy	3393	GAACGCTGTCGCTTCTGGCTGAGACTTCAAGAGC	3429
Db	202287	GATATATATGTCACATCTGGGAGGAGCTTAAAGAGC	202323

RESULT 15  
US-09-949-016-14581/C

; Sequence 14581, Application US/09949016  
; Patent No. 6812339

```

; GENERAL INFORMATION:
;
; APPLICANT: VENTER, J. Craig et al.
;
; TITLE OF INVENTION:

```

FILE REFERRED TO 001107

```

; FILE REFERENCE: CL00130/
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

```

PRIOR FILING DATE: 2000-04-19  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.00

```

```

; SEQ ID NO 14581
; LENGTH: 64309

```

```

; TYPE: DNA
; ORGANISM: Human

```

```
; FEATURE:
; NAME/KEY: misc_feature
```

```

; LOCATION: (1)...(64309)
; OTHER INFORMATION: n = A,T,C or G

```

US-09-949-016-14581

Query Match	1.3%;	Score 50.8;	DB 4;	Length 64309;
-------------	-------	-------------	-------	---------------

Best Local Similarity 58.7%; Pred.No. 0.0059;  
Matches 88; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

[illegible]

Search completed: July 19, 2005, 01:47:59  
Job time : 645 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 16:47:53 ; Search time 2336 Seconds  
(without alignments)  
10958.514 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atgagagagagaacaatcgct.....tcgaggsgctgctgctctaa 4035

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7173243 seqs, 3172129809 residues

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US10J\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4035	100.0	4035	US-10-124-498-5	Sequence 5, Appli
2	4035	100.0	4035	US-10-066-521-5	Sequence 5, Appli
3	3166.2	78.5	3885	US-10-860-761-3	Sequence 3, Appli
4	3166.2	78.5	3900	US-10-399-443-23	Sequence 23, Appli
5	3166.2	78.5	3900	US-10-677-943-23	Sequence 23, Appli
6	3139	77.8	3830	US-10-216-645-3	Sequence 3, Appli
7	3133.6	77.7	3489	US-10-416-642-3	Sequence 3, Appli

8	3133.6	77.7	3926	US-10-216-645-1	Sequence 1, Appli
9	2900.6	71.9	3226	US-10-092-900A-347	Sequence 347, App
10	1278	31.7	3447	US-10-399-443-5	Sequence 5, Appli
11	1278	31.7	3447	US-10-677-943-5	Sequence 5, Appli
12	1152.8	28.6	1157	US-10-399-443-1	Sequence 1, Appli
13	1152.8	28.6	1157	US-10-677-943-1	Sequence 1, Appli
14	756.6	18.8	1075	US-10-399-443-3	Sequence 3, Appli
15	756.6	18.8	1075	US-10-677-943-3	Sequence 3, Appli
16	729.4	18.1	3521	US-10-794-442-2	Sequence 2, Appli
17	438	10.9	2775	US-10-357-820-51	Sequence 51, Appli
18	424	10.5	2985	US-09-965-621-15	Sequence 15, Appli
19	424	10.5	2985	US-10-407-866-15	Sequence 15, Appli
20	424	10.5	2985	US-10-781-294-15	Sequence 15, Appli
21	422.4	10.5	2982	US-10-124-498-25	Sequence 25, Appli
22	422.4	10.5	2982	US-10-066-521-25	Sequence 25, Appli
23	422.4	10.5	3190	US-10-054-749-1079	Sequence 1079, Ap
24	422.4	10.5	3368	US-10-124-498-23	Sequence 23, Appli
25	422.4	10.5	3368	US-10-066-521-23	Sequence 23, Appli
26	419.2	10.4	2575	US-09-848-035-7	Sequence 7, Appli
27	419.2	10.4	2575	US-09-986-224-7	Sequence 7, Appli
28	419.2	10.4	2883	US-10-161-493-3	Sequence 3, Appli
29	419.2	10.4	3172	US-10-398-037-74	Sequence 74, Appli
30	409	10.1	3186	US-10-124-498-17	Sequence 17, Appli
31	409	10.1	3186	US-10-066-521-17	Sequence 17, Appli
32	404.4	10.0	3108	US-09-965-621-23	Sequence 23, Appli
33	404.4	10.0	3108	US-10-407-866-23	Sequence 23, Appli
34	404.4	10.0	3108	US-10-781-294-23	Sequence 23, Appli
35	404.4	10.0	3218	US-10-407-866-67	Sequence 67, Appli
36	392.6	9.7	3189	US-10-239-653-10	Sequence 10, Appli
37	392.6	9.7	4944	US-10-499-353A-429	Sequence 429, App
38	390.6	9.7	3540	US-10-794-442-7	Sequence 7, Appli
39	381.4	9.5	3099	US-10-127-516-3	Sequence 3, Appli
40	381.4	9.5	3099	US-10-027-629-3	Sequence 3, Appli
41	381.4	9.5	3099	US-10-132-567-3	Sequence 3, Appli
42	381.4	9.5	3431	US-10-127-516-1	Sequence 1, Appli
43	381.4	9.5	3431	US-10-027-629-1	Sequence 1, Appli
44	381.4	9.5	3431	US-10-132-567-1	Sequence 1, Appli
45	368.4	9.1	2763	US-10-250-615-39	Sequence 39, Appli

#### ALIGNMENTS

RESULT 1  
US-10-124-498-5  
; Sequence 5, Application US/10124498  
; Publication No. US20030017983A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; APPLICANT: Wang, Weiye  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
; FILE REFERENCE: 07334-367001  
; CURRENT APPLICATION NUMBER: US/10/124,498  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 10/066,521  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/318,645  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/265,231  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 4035  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(4032)  
US-10-124-498-5

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGAGCAAAATCGCTCACTTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60  
DB 1 ATGAGAGAGCAAAATCGCTCACTTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60

QY 61 CTAGCAAGAGAAATTTCAAGATTCAAGAAATTAATCAAGAAATCTTCAGAAATG 120  
DB 61 CTAGCAAGAGAAATTTCAAGATTCAAGAAATTAATCAAGAAATCTTCAGAAATG 120

QY 61 ACCATATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGGCACTC 180  
DB 61 CTAGCAAGAGAAATTTCAAGATTCAAGAAATTAATCAAGAAATCTTCAGAAATG 120

QY 121 ACCATATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGGCACTC 180  
DB 121 ACCATATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGGCACTC 180

QY 181 CTCTTGATGATATTAATGAGATCGTGGCTGGGCTACGTCATTAAGACATTTTGA 240  
DB 181 CTCTTGATGATATTAATGAGATCGTGGCTGGGCTACGTCATTAAGACATTTTGA 240

QY 241 AACATGAACCTGCGCAACCTCTCGAGAAAGCACGGGATGACATGAAAAATTTCAAA 300  
DB 241 AACATGAACCTGCGCAACCTCTCGAGAAAGCACGGGATGACATGAAAAATTTCAAA 300

QY 301 GCTATGAAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAAATTTCAAACT 360  
DB 301 GCTATGAAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAAATTTCAAACT 360

QY 361 ATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAAACAATGAGGTGACA 420  
DB 361 ATGGAACAAGAGTGCCACAGCAGCAGAGACAGAAACAAGAAACAATGAGGTGACA 420

QY 421 TGGGACTCAAGAGTCAAGTGAATGACAAATTCGCTGAGAGAGAGATGACGTGATG 480  
DB 421 TGGGACTCAAGAGTCAAGTGAATGACAAATTCGCTGAGAGAGAGATGACGTGATG 480

QY 481 TTTGAAAAACACTGCTGCTGACAGTGGCGGAAATGCAAAAGTTGGCTGTTTGAATCA 540  
DB 481 TTTGAAAAACACTGCTGCTGACAGTGGCGGAAATGCAAAAGTTGGCTGTTTGAATCA 540

QY 541 GACCGGTGGGCTTCGGGCTTCGCAAGGTGTTCTGCAAGAAAGTCAGAAATGGGAAA 600  
DB 541 GACCGGTGGGCTTCGGGCTTCGCAAGGTGTTCTGCAAGAAAGTCAGAAATGGGAAA 600

QY 601 TCGGCTCAGCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 601 TCGGCTCAGCCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 661 TTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATGACGGGAAAGAGAGAGAGAGTGC 720  
DB 661 TTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATGACGGGAAAGAGAGAGAGAGTGC 720

QY 721 ACAAGATTCATCTCCAGGAGTGCCAGACTCCAGGCTCCGTTGACGAGATCATGTC 780  
DB 721 ACAAGATTCATCTCCAGGAGTGCCAGACTCCAGGCTCCGTTGACGAGATCATGTC 780

QY 781 CGACCAAGAAAGGCTGTTTATCATGATGAGGTTTCAATGACCTGCTCTCTGCTGAC 840  
DB 781 CGACCAAGAAAGGCTGTTTATCATGATGAGGTTTCAATGACCTGCTCTCTGCTGAC 840

QY 841 AATGACAAAGAGCTCTGCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
DB 841 AATGACAAAGAGCTCTGCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

QY 901 AGTCTGCTGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 901 AGTCTGCTGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

QY 961 GGCACAGATC 1020  
DB 961 GGCACAGATC 1020

QY 1021 TCCGGGAAACAAGATCACTTGTCTTGAAGCGGGATGAGAGATCAAGAGACA 1080

DB 1021 TCCGGGAAACAAGATCACTTGTCTTGAAGCGGGATGAGAGATCAAGAGACA 1080

QY 1081 CAAGGTTGGCTGCTGATGAGACACCGTGAAGTCTCGACAGTGCAGGTGCCGCC 1140  
DB 1081 CAAGGTTGGCTGCTGATGAGACACCGTGAAGTCTCGACAGTGCAGGTGCCGCC 1140

QY 1141 GTGGGCTCTCATCTGAGTGGGCTGCTGAGTGGGAGAGAGAGAGAGAGAGAGAG 1200  
DB 1141 GTGGGCTCTCATCTGAGTGGGCTGCTGAGTGGGAGAGAGAGAGAGAGAGAGAG 1200

QY 1201 CCTTCAACCAAGAGCTCAGAGGCTGCAAGCGCTTTTGTGTTCAATGACCTCA 1260  
DB 1201 CCTTCAACCAAGAGCTCAGAGGCTGCAAGCGCTTTTGTGTTCAATGACCTCA 1260

QY 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
DB 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320

QY 1321 CGTATGCTGAG 1380  
DB 1321 CGTATGCTGAG 1380

QY 1381 GTTCAAG 1440  
DB 1381 GTTCAAG 1440

QY 1441 GACAGCCTGCTGAG 1500  
DB 1441 GACAGCCTGCTGAG 1500

QY 1501 GCTTGTACTAG 1560  
DB 1501 GCTTGTACTAG 1560

QY 1561 GAG 1620  
DB 1561 GAG 1620

QY 1621 ATGAG 1680  
DB 1621 ATGAG 1680

QY 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
DB 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

QY 1741 GGTGAG 1800  
DB 1741 GGTGAG 1800

QY 1801 GAGAGCTGAG 1860  
DB 1801 GAGAGCTGAG 1860

QY 1861 CGGATTAACAAG 1920  
DB 1861 CGGATTAACAAG 1920

QY 1921 TTGCGGAAAAATTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
DB 1921 TTGCGGAAAAATTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980

QY 1981 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
DB 1981 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040

QY 2041 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
DB 2041 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100

QY 2101 CTGACAGATC 2160

Db 2101 CTGACGAGCGGGCCATGAAAGACCTGTGTGCGAAAGCTGAGGATCCCACTGCAAGATTA 2160  
Qy 2161 CAGACCTGATGTTTAAAGAAATGCAAGATTAACCTGTGTGAGAGCACTCTGAGAAATC 2220  
Db 2161 CAGACCTGATGTTTAAAGAAATGCAAGATTAACCTGTGTGAGAGCACTCTGAGAAATC 2220  
Qy 2221 GTCATGGCCCAACCGTAACTTAAGATCCCTCAACTGGGAGGACCACTGAAAGAAAG 2280  
Db 2221 GTCATGGCCCAACCGTAACTTAAGATCCCTCAACTGGGAGGACCACTGAAAGAAAG 2280  
Qy 2281 GATGTAAAGATGCGGTGTGAAAGCTTTAAACAACCAAAATTTTGTGTGAGTCTTTGAGG 2340  
Db 2281 GATGTAAAGATGCGGTGTGAAAGCTTTAAACAACCAAAATTTTGTGTGAGTCTTTGAGG 2340  
Qy 2341 CTGGAATGCTGTGATGATGACCAATGCTGTAACTGGAAGATCTCCAAATCTTTACGACC 2400  
Db 2341 CTGGAATGCTGTGATGATGACCAATGCTGTAACTGGAAGATCTCCAAATCTTTACGACC 2400  
Qy 2401 TCCCCAGCGCTGAATCTCTGAGCGTGGCAGGAAACAAAGGTGACAGACCAAGGATTAATG 2460  
Db 2401 TCCCCAGCGCTGAATCTCTGAGCGTGGCAGGAAACAAAGGTGACAGACCAAGGATTAATG 2460  
Qy 2461 CCTCTCAGTATGCTCTTGAAGATCTCCAGTGCCTGTCAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTATGCTCTTGAAGATCTCCAGTGCCTGTCAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGGCATCAAGCCACGGGTTGCCAGATCTGGCTCAGCCCTGTGACGACCGAGAC 2580  
Db 2521 TGTGGCATCAAGCCACGGGTTGCCAGATCTGGCTCAGCCCTGTGACGACCGAGAC 2580  
Qy 2581 TTGACACACCTGTGCTATCCAAACAAGCGCTGGGAAACGAAGGTAAATCTACGTGT 2640  
Db 2581 TTGACACACCTGTGCTATCCAAACAAGCGCTGGGAAACGAAGGTAAATCTACGTGT 2640  
Qy 2641 CGATCCATGAGGCTTCCCACTGATGTGACAGAGCTGATCTGAATCAAGTCCACCTG 2700  
Db 2641 CGATCCATGAGGCTTCCCACTGATGTGACAGAGCTGATCTGAATCAAGTCCACCTG 2700  
Qy 2701 GACACGGCTGCGTGTGTTTCTTGCACTTGGCTTAATGAGTAACTCATGCTGACGAC 2760  
Db 2701 GACACGGCTGCGTGTGTTTCTTGCACTTGGCTTAATGAGTAACTCATGCTGACGAC 2760  
Qy 2761 CTGACCGTTAGCATTAACCTGTGGAAGCAATGACGTGAAGCTTCTGCGAGGTCAATG 2820  
Db 2761 CTGACCGTTAGCATTAACCTGTGGAAGCAATGACGTGAAGCTTCTGCGAGGTCAATG 2820  
Qy 2821 AGAGAACCATCTTGTCACTCCAGACCTGAGAGTGTAAAGTGTCACTCAACCGCCG 2880  
Db 2821 AGAGAACCATCTTGTCACTCCAGACCTGAGAGTGTAAAGTGTCACTCAACCGCCG 2880  
Qy 2881 TGTGTGAGAGTCTGTCTGTGTGATCTTGAAGACAGACCTGAAGAGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTCTGTCTGTGTGATCTTGAAGACAGACCTGAAGAGCTGTGATCTC 2940  
Qy 2941 ACGGACATATGCGGTGGTGAACGCTGGGTGTGCTGCTGTGAGAGGACGTAAGCAAAAG 3000  
Db 2941 ACGGACATATGCGGTGGTGAACGCTGGGTGTGCTGCTGTGAGAGGACGTAAGCAAAAG 3000  
Qy 3001 AACAGTGTCTGACAGACTCGGGTTGAAGCATGTGACCTGATCTTGAATGCTGTGAG 3060  
Db 3001 AACAGTGTCTGACAGACTCGGGTTGAAGCATGTGACCTGATCTTGAATGCTGTGAG 3060  
Qy 3061 GCACCTCTCTTGGCCCTTTCTGCAACCGGCATCTGACCAAGTCTAAACCTGTGTGAGAT 3120  
Db 3061 GCACCTCTCTTGGCCCTTTCTGCAACCGGCATCTGACCAAGTCTAAACCTGTGTGAGAT 3120  
Qy 3121 AACCTCAGTCCCAAGAAATGATGAAGCTGTGTGGCTTTGGCTGTGCTCCCAAGCTTAAC 3180  
Db 3121 AACCTCAGTCCCAAGAAATGATGAAGCTGTGTGGCTTTGGCTGTGCTCCCAAGCTTAAC 3180  
Qy 3181 TTACAGATTAATGGGCTGTGAAATGAGCATGACCTGTGCAAAATGAAGAGCTGTGAG 3240  
Db 3181 TTACAGATTAATGGGCTGTGAAATGAGCATGACCTGTGCAAAATGAAGAGCTGTGAG 3240

Qy 3241 GAAGTGCAGTACTCAAGCCCGAGTGTGTAATGAACGCTAGTGGCATTTCTTTGATGA 3300  
Db 3241 GAAGTGCAGTACTCAAGCCCGAGTGTGTAATGAACGCTAGTGGCATTTCTTTGATGA 3300  
Qy 3301 GATGACCGACCAAAATAGACTTAATCTTCCGCGCTCCCTGAAAGCGCGGCAATGGCATGT 3360  
Db 3301 GATGACCGACCAAAATAGACTTAATCTTCCGCGCTCCCTGAAAGCGCGGCAATGGCATGT 3360  
Qy 3361 GCTTGTGTGTGGGATGAACCCAGAGCAAAAGCGTGTGTGCTTCTGTGAGAGAC 3420  
Db 3361 GCTTGTGTGTGGGATGAACCCAGAGCAAAAGCGTGTGTGCTTCTGTGAGAGAC 3420  
Qy 3421 TTCAAAGACAGTAAACGATTTTCCAAAGTCTCTGCGCTGGCCAGCGCAAAATGAGATCC 3480  
Db 3421 TTCAAAGACAGTAAACGATTTTCCAAAGTCTCTGCGCTGGCCAGCGCAAAATGAGATCC 3480  
Qy 3481 CAGAGAGTTGACAAAGTGTGACAGAGCTCCCGCAACCATGAGGACGAGCAAGCAACAA 3540  
Db 3481 CAGAGAGTTGACAAAGTGTGACAGAGCTCCCGCAACCATGAGGACGAGCAAGCAACAA 3540  
Qy 3541 CAAGTAAATGTTGATGTTGATTAATCCGAGCTGTGTGAAATGCTGAGCTCGAA 3600  
Db 3541 CAAGTAAATGTTGATGTTGATTAATCCGAGCTGTGTGAAATGCTGAGCTCGAA 3600  
Qy 3601 GGGCTTGTATCCAAACAGTGTGATCAATGACCAAGAGGTATGGCTGTGCTGAGGAG 3660  
Db 3601 GGGCTTGTATCCAAACAGTGTGATCAATGACCAAGAGGTATGGCTGTGCTGAGGAG 3660  
Qy 3661 GAGCTGAGCTGAGGAGCTGTGTGTCAAACAGTGTGATGACCAAGAGGTGTGTCTGCT 3720  
Db 3661 GAGCTGAGCTGAGGAGCTGTGTGTCAAACAGTGTGATGACCAAGAGGTGTGTCTGCT 3720  
Qy 3721 CACTGGAGCGGCTGTGGCTTTAGGGGCTGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780  
Db 3721 CACTGGAGCGGCTGTGGCTTTAGGGGCTGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780  
Qy 3781 GTGTCTGTGTCACTGGGAGCGGCTGTGGCTGTGGGAGCTGTGTCTTAAACAGTGTGAT 3840  
Db 3781 GTGTCTGTGTCACTGGGAGCGGCTGTGGCTGTGGGAGCTGTGTCTTAAACAGTGTGAT 3840  
Qy 3841 CACAGCGGTGTGGCTGTGTGTCATGGGAGCGGCTGTGGGAGCTGTGTCTTAAACAGT 3900  
Db 3841 CACAGCGGTGTGGCTGTGTGTCATGGGAGCGGCTGTGGGAGCTGTGTCTTAAACAGT 3900  
Qy 3901 GCTATGACCAAGCGGTGTGTCTGTGTCACTGGGAGCGGCTGTGGGAGCTGTGTG 3960  
Db 3901 GCTATGACCAAGCGGTGTGTCTGTGTCACTGGGAGCGGCTGTGGGAGCTGTGTG 3960  
Qy 3961 TCCAACAGTGTGATGACCAAGCGGTGTGTCTGTGTCACTGGGAGCGGCTGTGGAG 4020  
Db 3961 TCCAACAGTGTGATGACCAAGCGGTGTGTCTGTGTCACTGGGAGCGGCTGTGGAG 4020  
Qy 4021 GGGCTGTGTCTTAA 4035  
Db 4021 GGGCTGTGTCTTAA 4035

RESULT 2  
US-10-066-521-5  
Sequence 5, Application US/10066521  
Publication No. US2003027757A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Wang, Weiye  
APPLICANT: Blacher, Maria  
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
FILE REFERENCE: 07334-334001  
CURRENT APPLICATION NUMBER: US/10/066, 521  
CURRENT FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: 60/318, 645  
PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/265,231  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 4035  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(4032)  
US-10-066-521-5

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGACAAATGCTCACTTTTCAGCTACGGGCTGCAATGCTGCTATGAG 60  
DB 1 ATGGAAGAGACAAATGCTCACTTTTCAGCTACGGGCTGCAATGCTGCTATGAG 60  
QY 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAATTAAGAAATTTCAAGAAATG 120  
DB 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAATTAAGAAATTTCAAGAAATG 120  
QY 121 ACCACATGCTCTATTCACAGATTGAAATCGAATGCAAGTGAATGCTGCACTC 180  
DB 121 ACCACATGCTCTATTCACAGATTGAAATCGAATGCAAGTGAATGCTGCACTC 180  
QY 181 CTCTTCATGATATTATGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 181 CTCTTCATGATATTATGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAATTTCA 300  
DB 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAATTTCA 300  
QY 301 GCTATGAAACAAGAGTGCACAGCAGCAGAGACAGAAATTTCAAGAAATTTCAAG 360  
DB 301 GCTATGAAACAAGAGTGCACAGCAGCAGAGACAGAAATTTCAAGAAATTTCAAG 360  
QY 361 ATGGAACAAGAGTGCACAGCAGCAGAGACAGAAATTTCAAGAAATTTCAAG 420  
DB 361 ATGGAACAAGAGTGCACAGCAGCAGAGACAGAAATTTCAAGAAATTTCAAG 420  
QY 421 TGGGCTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGCTGCTGCTG 480  
DB 421 TGGGCTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGCTGCTGCTG 480  
QY 481 TTTGAAAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 481 TTTGAAAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 541 GACCGGTGGGCTTCGCGCTGCGACGGTGTCTGCAAGAAATTTCAAGAAATTT 600  
DB 541 GACCGGTGGGCTTCGCGCTGCGACGGTGTCTGCAAGAAATTTCAAGAAATTT 600  
QY 601 TCGGCTAGCCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 601 TCGGCTAGCCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 661 TTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACAGCGAGAGAGAGAGAG 720  
DB 661 TTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACAGCGAGAGAGAGAGAG 720  
QY 721 ACAGAGTTCATCTCAGGGAAGTGGCCAGACTCCAGGCTCGGTTACCGAGATCA 780  
DB 721 ACAGAGTTCATCTCAGGGAAGTGGCCAGACTCCAGGCTCGGTTACCGAGATCA 780  
QY 781 CGACCAAGAAAGGCTGTTGATCAATGACGGTTCGATGACCTGGGCTGCTGCTG 840  
DB 781 CGACCAAGAAAGGCTGTTGATCAATGACGGTTCGATGACCTGGGCTGCTGCTG 840

QY 841 AATGACAAAGCTCTGCAAAAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
DB 841 AATGACAAAGAGCTCTGCAAAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
QY 901 AGTCTGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
DB 901 AGTCTGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
QY 961 GGCACAGAGAGAGCTCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
DB 961 GGCACAGAGAGAGCTCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1021 TCCGAGAGAAAGAAATCACTTCTGCTGAGCGAGATGCTGAGATCAAGAGACA 1080  
DB 1021 TCCGAGAGAAAGAAATCACTTCTGAGCGAGATGCTGAGATCAAGAGACA 1080  
QY 1081 CAAGGTTGCTGCTGATCAATGACACCGTGAAGTCTGCAAGAGAGAGAGAGAG 1140  
DB 1081 CAAGGTTGCTGATCAATGACACCGTGAAGTCTGCAAGAGAGAGAGAGAG 1140  
QY 1141 GTGGGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
DB 1141 GTGGGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1201 CCTTCAACCAAGAGCTCAAGAGCTGCAAGAGCTGCTGCTGCTGCTGCTGCTG 1260  
DB 1201 CCTTCAACCAAGAGCTCAAGAGCTGCAAGAGCTGCTGCTGCTGCTGCTGCTG 1260  
QY 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
DB 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
QY 1321 CGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
DB 1321 CGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1381 GTTCAAGAGAGCTGCGGAGAGCTGAGCTCGTCTGCTGCTGCTGCTGCTGCTG 1440  
DB 1381 GTTCAAGAGAGCTGCGGAGAGCTGAGCTCGTCTGCTGCTGCTGCTGCTGCTG 1440  
QY 1441 GACAGCACTGCTGAGAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1441 GACAGCACTGCTGAGAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 GCTTGTACTACGCTGCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
DB 1501 GCTTGTACTACGCTGCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
QY 1561 GAGAGACAAAGAGAGCTCATGAGAGCTTAAACAGGCAAGCTTCACTGCTGCTG 1620  
DB 1561 GAGAGACAAAGAGAGCTCATGAGAGCTTAAACAGGCAAGCTTCACTGCTGCTG 1620  
QY 1621 ATGAGAGGTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
DB 1621 ATGAGAGGTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
QY 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
DB 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
QY 1741 GGTGAGAGAGCTTAATGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 1741 GGTGAGAGAGCTTAATGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
QY 1801 GAGACTCAAGCAAGAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860  
DB 1801 GAGACTCAAGCAAGAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860  
QY 1861 CCGATTAAACAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
DB 1861 CCGATTAAACAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
QY 1921 TTGCGAGAAATTCGGGTGAGATGTCAAAGAGAGATCTTCCAAAGAGATGAGTCCG 1980



Db 1921 TTGGGAAAAATTGGGGTGGATGTCAAAGGGATCTTCCCAAGAGATGAGTCCCTGAGGCA 1980  
Qy 1981 TGTCTGTGAGTCCCTCTATGATGCGGGATAAGACCTTATTGAGAGAGAGTGGAAAT 2040  
Db 1981 TGTCTGTGAGTCCCTCTATGATGCGGGATAAGACCTTATTGAGAGAGAGTGGAAAT 2040  
Qy 2041 TTCTGCTCCATGCTTGGCAACCACTGCGGAGCTGGAACCTGGGAGAGACATC 2100  
Db 2041 TTCTGCTCCATGCTTGGCAACCACTGCGGAGCTGGAACCTGGGAGAGACATC 2100  
Qy 2101 CTGACAGAGCGGGCCATGAAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGGCCATGAAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTTGAAGATGACAGATTAACCTGTGTGCAAGACCTTGAAGATC 2220  
Db 2161 CAGACCTGATGTTTGAAGATGACAGATTAACCTGTGTGCAAGACCTTGAAGATC 2220  
Qy 2221 GTGATGGGCAACCGTAACCTAGATCCCTCACTTGGGAGGACCACTGAAAGAGAG 2280  
Db 2221 GTGATGGGCAACCGTAACCTAGATCCCTCACTTGGGAGGACCACTGAAAGAGAG 2280  
Qy 2281 GATGTAAGGATGGCTGTGAGACCTTAAACACCCAAATGTTGTGAGTCTTGAAG 2340  
Db 2281 GATGTAAGGATGGCTGTGAGACCTTAAACACCCAAATGTTGTGAGTCTTGAAG 2340  
Qy 2341 CTGATGTGTGTGATGATGACCATGCTGTGTAACCTGAAGATCTCCAAATCTTACGAC 2400  
Db 2341 CTGATGTGTGTGATGATGACCATGCTGTGTAACCTGAAGATCTCCAAATCTTACGAC 2400  
Qy 2401 TCCCCAGCTGGAATCTGTAGCCTGGAGGAAACAAGTGAACAACAGGAGATTAATG 2460  
Db 2401 TCCCCAGCTGGAATCTGTAGCCTGGAGGAAACAAGTGAACAACAGGAGATTAATG 2460  
Qy 2461 CCTCTCAGTATGCTTGTGAGAGTCTCCAGTGTGCGCTGCAAGACCTGATATCTGAGAG 2520  
Db 2461 CCTCTCAGTATGCTTGTGAGAGTCTCCAGTGTGCGCTGCAAGACCTGATATCTGAGAG 2520  
Qy 2521 TGTGGCATGACAGCAACCGGCTTGCAGAGTGTGCGCTGCAAGACCTGATATCTGAGAG 2580  
Db 2521 TGTGGCATGACAGCAACCGGCTTGCAGAGTGTGCGCTGCAAGACCTGATATCTGAGAG 2580  
Qy 2581 TTGACACACCTGTGCTATCCACACACAGCTGGGGAAGAGGTGAATTTACTGTGT 2640  
Db 2581 TTGACACACCTGTGCTATCCACACACAGCTGGGGAAGAGGTGAATTTACTGTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATCAGTGCACCTG 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATCAGTGCACCTG 2700  
Qy 2701 GACACGGCTGTGTGTTTCTTGTGCACTTGTGCTTATGAGTACTCATGCTGACGAC 2760  
Db 2701 GACACGGCTGTGTGTTTCTTGTGCACTTGTGCTTATGAGTACTCATGCTGACGAC 2760  
Qy 2761 CTGAGCTTATGATGAACTTGTGAAACAATGCGTGAAGCTTGTGTGAGGATCATG 2820  
Db 2761 CTGAGCTTATGATGAACTTGTGAAACAATGCGTGAAGCTTGTGTGAGGATCATG 2820  
Qy 2821 AGAGAACCATCTTGTATCTCCAGGACCTGAGGTTGTAAAGTGTATCTCACTGACGCGCG 2880  
Db 2821 AGAGAACCATCTTGTATCTCCAGGACCTGAGGTTGTAAAGTGTATCTCACTGACGCGCG 2880  
Qy 2881 TGTCTGTGAGTCTGT 2940  
Db 2881 TGTCTGTGAGTCTGT 2940  
Qy 2941 ACGGACATGCTTGT 3000  
Db 2941 ACGGACATGCTTGT 3000  
Qy 3001 AACAGTGTCTGACAGAGCTGCGGTTGAAGGCAATGAGCTGATCTTGAATGTGTGTGTG 3060  
Db 3001 AACAGTGTCTGACAGAGCTGCGGTTGAAGGCAATGAGCTGATCTTGAATGTGTGTGTG 3060

Db 3001 AACAGTGTCTGACAGAGCTGCGGTTGAAGGCAATGAGCTGATCTTGAATGTGTGTGTG 3060  
Qy 3061 GCATCTCTCTTGGCCCTTCTTCTGCAACGGGATCTGACAGCTTAAACCTGTGTGAGAT 3120  
Db 3061 GCATCTCTCTTGGCCCTTCTTCTGCAACGGGATCTGACAGCTTAAACCTGTGTGAGAT 3120  
Qy 3121 AACTTCAGTCCCAAGATGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180  
Db 3121 AACTTCAGTCCCAAGATGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180  
Qy 3181 TTACAGATTAATGTGTGTGAAATGCAAGTACCTGTGTGCAATTAAGAGAGCTGTGAG 3240  
Db 3181 TTACAGATTAATGTGTGTGAAATGCAAGTACCTGTGTGCAATTAAGAGAGCTGTGAG 3240  
Qy 3241 GAGTGTGAGTCTTAAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300  
Db 3241 GAGTGTGAGTCTTAAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300  
Qy 3301 GATGACGACCAAAATGAGACTTACTTCCGGCTCCCTGAAGCCGGGCAATGGCATGT 3360  
Db 3301 GATGACGACCAAAATGAGACTTACTTCCGGCTCCCTGAAGCCGGGCAATGGCATGT 3360  
Qy 3361 GCTTGTGTGTGTGTGTGAAACCAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
Db 3361 GCTTGTGTGTGTGTGTGAAACCAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
Qy 3421 TTCAAGAGAGTACAGATTTGTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480  
Db 3421 TTCAAGAGAGTACAGATTTGTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480  
Qy 3481 CAGAGAGTTGAACAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
Db 3481 CAGAGAGTTGAACAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
Qy 3541 CAGAGAGTTGAACAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
Db 3541 CAGAGAGTTGAACAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
Qy 3601 GGGCTTGT 3660  
Db 3601 GGGCTTGT 3660  
Qy 3661 GAGCTGT 3720  
Db 3661 GAGCTGT 3720  
Qy 3721 CACTGT 3780  
Db 3721 CACTGT 3780  
Qy 3781 GTGTCTGT 3840  
Db 3781 GTGTCTGT 3840  
Qy 3841 CACAGCGT 3900  
Db 3841 CACAGCGT 3900  
Qy 3901 GCTGT 3960  
Db 3901 GCTGT 3960  
Qy 3961 TCCAAAGT 4020  
Db 3961 TCCAAAGT 4020  
Qy 4021 GGGCTGT 4080  
Db 4021 GGGCTGT 4080

Sequence 3, Application US/10860761  
Publication No. US2004024875A1  
GENERAL INFORMATION:  
APPLICANT: MYETH  
TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS  
FILE REFERENCE: AM101318  
CURRENT APPLICATION NUMBER: US/10/860,761  
CURRENT FILING DATE: 2004-06-03  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO: 3  
LENGTH: 3885  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(3603)  
US-10-860-761-3

Query Match 78.5%; Score 316.2; DB 20; Length 3885;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

QY 1 ATGGAAGGAGCAAAATCGCTCACTTTTCCAGCTACGGGCTGCAATGGTGTCTATGAG 60  
DB 154 ATGGAAGGAGCAAAATCGCTCACTTTTCCAGCTACGGGCTGCAATGGTGTCTATGAG 213  
QY 61 CTAGCAAGGAAGAAATTCAGCATTCAGGAATTAAGAAAGAAATCTTCAGAAATCG 120  
DB 214 CTAGCAAGGAAGAAATTCAGCATTCAGGAATTAAGAAAGAAATCTTCAGAAATCG 273  
QY 121 ACCAATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGCACTC 180  
DB 274 ACCAATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGCACTC 333  
QY 181 CTCTTGCAATGATTAATGAGCATCGTGGCTGAGTCAATGATGATCTTTGAA 240  
DB 334 CTCTTGCAATGATTAATGAGCATCGTGGCTGAGTCAATGATGATCTTTGAA 393  
QY 241 AACATGAACCTGCAACCTCTCGGAGAAAGGACGGGATGACATGAAAA----- 289  
DB 394 AACATGAACCTGCAACCTCTCGGAGAAAGGACGGGATGACATGAAAAAGCATTCACCA 453  
QY 290 ----- 289  
DB 454 GAAGATCTGAAGCAAGATGACGACCAAGGACCAAGAAAGTGCAGAAAT 513  
QY 290 -----AAATTTCA 297  
DB 514 TCACAAAGCTGTGCAACAAGATAGTGCACAGCTGACAGACAAAAGAAACAGAAATTTCA 573  
QY 298 CAAGCTATGGAACAAGAGGTGCCACAGCAGACAGACAGAAAGAAATTTCAAA 357  
DB 574 CAAGCTATGGAACAAGAGGTGCCACAGCAGACAGACAGAAAGAAATTTCAAA 633  
QY 358 GCTATGGAACAAGAGGTGCCACAGCAGACAGACAGAAAGAAAGATGAGAGTAC 417  
DB 634 GCTATGGAACAAGAGGTGCCACAGCAGACAGACAGAAAGAAAGATGAGAGTAC 693  
QY 418 ACATGGGACCTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGATGCTGT 477  
DB 694 ACATGGGACCTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGATGCTGT 753  
QY 478 AGTTTGAACCACTGCTGACCTGAGCGGAAATGCAAGTGGGCTGGCTTTTAT 537  
DB 754 AGTTTGAACCACTGCTGACCTGAGCGGAAATGCAAGTGGGCTGGCTTTTAT 813  
QY 538 TCAGACCGGTGGGCTTCGCGCTCGACAGGTGTTCTGCAAGGAAATGAGAAATTTGG 597  
DB 814 TCAGACCGGTGGGCTTCGCGCTCGACAGGTGTTCTGCAAGGAAATGAGAAATTTGG 873  
QY 598 AAATCGGCTTAAGCCAGGAAGATGCTGTGCTGGGCGCAAGGTGACCTTACAGGGA 657  
|||||

DB 874 AAATCGGCTTAAGCCAGGAAGATGCTGTGCTGGGCGCAAGGTGACCTTACAGGGA 933  
QY ATGTTCTCTACAGTCTTCTCTCCCGTTAGAGAGATGACGCGAAGAGAGAGAGT 717  
DB 934 ATGTTCTCTACAGTCTTCTCTCCCGTTAGAGAGATGACGCGAAGAGAGAGAGT 993  
QY 718 GTCCAGAGTTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGAGATCATG 777  
DB 994 GTCCAGAGTTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGAGATCATG 1053  
QY 778 TCCCGACCAAGAAAGCTGTTTATCATATGACGCTTTCATGACCTGGGCTCTGTCTC 837  
DB 1054 TCCCGACCAAGAAAGCTGTTTATCATATGACGCTTTCATGACCTGGGCTCTGTCTC 1113  
QY 838 AACATGACCAAAAGCTCTGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897  
DB 1114 AACATGACCAAAAGCTCTGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173  
QY 898 CGAGTCTGTGAGAAAGTCTGCTCTGAGTCTTCTGATCGTCAACGCTCAGAGAC 957  
DB 1174 CGAGTCTGTGAGAAAGTCTGCTCTGAGTCTTCTGATCGTCAACGCTCAGAGAC 1233  
QY 958 GTGGGCAAGAGAACTCAAGTCAAGAGTGTGTCTCCGTTTACCTTTAGTTAGAGA 1017  
DB 1234 GTGGGCAAGAGAACTCAAGTCAAGAGTGTGTCTCCGTTTACCTTTAGTTAGAGA 1293  
QY 1018 ATCTCCGGGGAAACAAAGATTCACATGCTCTTGAAGCGGAGATTTGTTAGATCAGAA 1077  
DB 1294 ATCTCCGGGGAAACAAAGATTCACATGCTCTTGAAGCGGAGATTTGTTAGATCAGAA 1353  
QY 1078 ACACAAAGGTTGCTGTGATCATGAACACGTTAGTGTCTGACACAGTGCAGAGTCCC 1137  
DB 1354 ACACAAAGGTTGCTGTGATCATGAACACGTTAGTGTGTCTGACACAGTGCAGAGTCCC 1413  
QY 1138 GCCGTGGCTCTCTCATCTGCGTGGCCCTGCAAGTGAAGAGAGAGAGAGAGAGTCC 1197  
DB 1414 GCCGTGGCTCTCTCATCTGCGTGGCCCTGCAAGTGAAGAGAGAGAGAGAGTCC 1473  
QY 1198 GCCCCTTCAACAAAGCTCACAGGCTGCAAGCCTTTTGTGTTCAATCAGCTCACC 1257  
DB 1474 GCCCCTTCAACAAAGCTCACAGGCTGCAAGCCTTTTGTGTTCAATCAGCTCACC 1533  
QY 1258 CTTGAGGCGTGTCCGCGCTGTCTCAATCTGAGAGAAAGATTTGCTGAAAGCGCTTC 1317  
DB 1534 CTTGAGGCGTGTCCGCGCTGTCTCAATCTGAGAGAAAGATTTGCTGAAAGCGCTTC 1593  
QY 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAATGATGTTTGAATGTGAGACCTC 1377  
DB 1594 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAATGATGTTTGAATGTGAGACCTC 1653  
QY 1378 ATGTTTCAAGGACTCGGAGAGTCTGAGCTCGTGTCTGTTTCAATGAACATCTTCTC 1437  
DB 1554 ATGTTTCAAGGACTCGGAGAGTCTGAGCTCGTGTCTGTTTCAATGAACATCTTCTC 1713  
QY 1438 CCAGACGCACTGTGAGAGATCTACACCTTCTTCACTCACTCACTCACTCACTTCTGT 1497  
DB 1714 CCAGACGCACTGTGAGAGATCTACACCTTCTTCACTCACTCACTCACTCACTTCTGT 1773  
QY 1498 GCCGCTTTGATACGTGTTAGAGAGGCTTGAATGAGGACGCTCTCTGCTGTAC 1557  
DB 1774 GCCGCTTTGATACGTGTTAGAGAGGCTTGAATGAGGACGCTCTCTGCTGTGTAC 1833  
QY 1558 GTTGAAGAAACAAAGAGTCTCATGAGCTTAAACAGGAGGCTTCAATATCACTGCTT 1617  
DB 1834 GTTGAAGAAACAAAGAGTCTCATGAGCTTAAACAGGAGGCTTCAATATCACTGCTT 1893  
QY 1618 TGAATGAAGGCTTCTTGTGTGCTGTGAGCGAAGACGTTAAGAGAGCACTGAGAGTCC 1677  
DB 1894 TGAATGAAGGCTTCTTGTGTGCTGTGAGCGAAGACGTTAAGAGAGCACTGAGAGTCC 1953  
QY 1678 CTGCTGGGCTGTCCGCTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTTCTCTG 1737  
DB 1954 CTGCTGGGCTGTCCGCTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTTCTCTG 2013  
|||||

1738 TTGGGTCAGAGCCTTAATGSCACACCCAGAGACACCTGAGAGCCTTCCACTGCTT 1797  
1739 TTTGGGTCAGAGCCTTAATGSCACACCCAGAGAGACCTTGGAGCCTTCCACTGCTT 2073  
1738 TTGAGACTCAAGACCAAGAGTTGTTGCTGGCATTAAGAGCTTCCAGAGAGTGTG 1857  
2074 TTGAGAGCTCAAGACCAAGAGTTGTTGCTGGCATTAAGAGCTTCCAGAGAGTGTG 2133  
1858 CTTCCGATTAAACAGAACCTGAGACTTGTATAGACTTCTTCTGCTCCAGAGCTGTCC 1917  
2134 CTTCCGATTAAACAGAACCTGAGACTTGTATAGACTTCTTCTGCTCCAGAGCTGTCC 2193  
1918 TATTTGCGGAAATTTGGGTGATGTCAGAGAGATTTTCCAGAGATGATGCTGTG 1977  
2194 TATTTGCGGAAATTTGGGTGATGTCAGAGAGATTTTCCAGAGATGATGCTGTG 2253  
1978 GCATGCTCTGTGGTCCCTCTATGATGCGGATTAAGACCTCATTTAGAGAGAGTGGAA 2037  
2254 GCATGCTCTGTGGTCCCTCTATGATGCGGATTAAGACCTCATTTAGAGAGAGTGGAA 2313  
2038 GATTTCTGCTCCATGCTTGGACCCACCCAGACTGCGGAGCTGAGCCTGGGAGAGCAG 2097  
2314 GATTTCTGCTCCATGCTTGGACCCACCCAGACTGCGGAGCTGAGCCTGGGAGAGCAG 2373  
2098 ATCTGACAGAGCGGCGCATGAGAGCCTGTGTGCCAGCTGAGGATCCACCTGCAAG 2157  
2374 ATCTGACAGAGCGGCGCATGAGAGCCTGTGTGCCAGCTGAGGATCCACCTGCAAG 2433  
2158 ATACAGACCTGTATGTTTGAATGACAGATTACCCCTGTGTGAGAGACCTGTGAGA 2217  
2434 ATACAGACCTGTATGTTTGAATGACAGATTACCCCTGTGTGAGAGACCTGTGAGA 2493  
2218 ATGCTCATGCGCAACCGTAACCTAAGATCCCTCAACTTGGAGAGACCCACCTGAAGAA 2277  
2494 ATGCTCATGCGCAACCGTAACCTAAGATCCCTCAACTTGGAGAGACCCACCTGAAGAA 2553  
2278 GAGAGTGTAGAGATGCGGTGTGAGAGCTTAAACCCAAATGTTTGTGAGAGTCTTG 2337  
2554 GAGAGTGTAGAGATGCGGTGTGAGAGCTTAAACCCAAATGTTTGTGAGAGTCTTG 2613  
2338 AGGCGGATGCTGTGATGAGAGCTTCCAGTCTGTACCTGAAGATCCCTCAACTTGG 2397  
2614 AGGCGGATGCTGTGATGAGAGCTTCCAGTCTGTACCTGAAGATCCCTCAACTTGG 2673  
2398 ACCCTCCCAAGCCTGAATCTCTGAGCCTGAGAGAAACAGAGTACAGACAGAGAGTA 2457  
2674 ACCCTCCCAAGCCTGAATCTCTGAGCCTGAGAGAAACAGAGTACAGACAGAGAGTA 2733  
2458 ATGCTCTCATGATGAGCTTGAAGTCTTCCAGTGTGCGCTGAGAGAGCTGATGAG 2517  
2734 AGGCTCTCATGATGAGCTTGAAGTCTTCCAGTGTGCGCTGAGAGAGCTGATGAG 2793  
2518 GACTGTGATGACAGAGCAACCGGTTGCCAGAGTCTGAGCTGAGCCCTGTGAGAGAGCG 2577  
2794 GACTGTGATGACAGAGCAACCGGTTGCCAGAGTCTGAGCTGAGCCCTGTGAGAGAGCG 2853  
2578 AGCTTGACACACTGTGCTATCAACAAACAGCTGGGAGAGAGAGTGTAAATCTACTG 2637  
2854 AGCTTGACACACTGTGCTATCAACAAACAGCTGGGAGAGAGAGTGTAAATCTACTG 2913  
2638 TGTGATTCATGAGGCTTCCCACTGTAGTCTGAGAGAGCTGATGAGAGAGTGTCCAC 2697  
2914 TGTGATTCATGAGGCTTCCCACTGTAGTCTGAGAGAGCTGATGAGAGAGTGTCCAC 2973  
2698 CTGAGACAGGCTGAGCTGTGCTTCTTGAAGCTTGGAGTAACTGAGAGCTGAGAG 2757  
2974 CTGAGACAGGCTGAGCTGTGCTTCTTGAAGCTTGGAGTAACTGAGAGCTGAGAG 3033  
2758 CACCTGAGCTTATGATGAACCTGTGAGAGAGAGAGTGTGAGAGCTTGTGAGAGT 2817  
3034 CACCTGAGCTTATGATGAACCTGTGAGAGAGAGAGTGTGAGAGCTTGTGAGAGT 3093

QY 2818 ATGAGAGAACCTTGTGATCTCCAGAGCCTGAGAGTGTGTAAGTGTATCTACCGCC 2877  
DB 3094 ATGAGAGAACCTTGTGATCTCCAGAGCCTGAGAGTGTGTAAGTGTATCTACCGCC 3153  
QY 2878 GCGTCTGTGAGAGCTGTCTGTGTGATCTGAGAGACAGACACTGAGAGCTGTGAT 2937  
DB 3154 GCGTCTGTGAGAGCTGTCTGTGTGATCTGAGAGACAGACACTGAGAGCTGTGAT 3213  
QY 2938 CTCAGGACATGCTGCTGAGTACGCTGGGTTGCTGCTGTGTGAGAGAGTGTGAT 2997  
DB 3214 CTCAGGACATGCTGCTGAGTACGCTGGGTTGCTGCTGTGTGAGAGAGTGTGAT 3273  
QY 2998 AAGACAGTGTGCTGAGAGCTGCTGGTGTGAGAGTGTGATCTGATGCTGTGAT 3057  
DB 3274 AAGACAGTGTGCTGAGAGCTGCTGGTGTGAGAGTGTGATCTGATGCTGTGAT 3333  
QY 3058 GAGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3117  
DB 3334 GAGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3393  
QY 3118 AATACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3177  
DB 3394 AATACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3453  
QY 3178 AACTTACAGATTAATGCTGTGAGATGAGAGTGTGCTGCTGCTGCTGCTGCTG 3237  
DB 3454 AACTTACAGATTAATGCTGTGAGATGAGAGTGTGCTGCTGCTGCTGCTGCTG 3513  
QY 3238 GAGAGAGTGCAGCTACTCAAGCCCGAGTGTGATGAGAGTGTGCTGCTGCTGCT 3297  
DB 3514 GAGAGAGTGCAGCTACTCAAGCCCGAGTGTGATGAGAGTGTGCTGCTGCTGCT 3573  
QY 3298 GAGAGTGCAGCTACTCAAGCCCGAGTGTGATGAGAGTGTGCTGCTGCTGCT 3312  
DB 3574 GAGAGTGCAGCTACTCAAGCCCGAGTGTGATGAGAGTGTGCTGCTGCTGCT 3588

RESULT 4  
US-10-399-443-23  
Sequence 23, Application US/10399443  
Publication No. US2004002869A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as Represented by the  
APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
APPLICANT: Health  
APPLICANT: Nelson, Lawrence M.  
APPLICANT: Tong, Zhi-Bin  
APPLICANT: Nelson, Lawrence  
APPLICANT: Zhi-Bin, Tong  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64785  
CURRENT APPLICATION NUMBER: US/10/399,443  
CURRENT FILING DATE: 2003-04-16  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 23  
LENGTH: 3900  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(3600)  
OTHER INFORMATION:  
US-10-399-443-23

Query Match 78.5%; Score 3166.2; DB 17; Length 3900;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

QY 1 ATGGAAGAGACAAATCGCTCTTTTCAGCTACGGGCTGCAATGCTCTATGAG 60  
Db 154 ATGGAAGAGACAAATCGCTCACCTTTTCAGCTACGGGCTGCAATGCTCTATGAG 213  
QY 61 CTGACCAAGGAAGATTTCAACATTCAGGAATTTACTAAAGAAATCTTCAGATCG 120  
Db 214 CTGACCAAGGAAGATTTCAACATTCAGGAATTTACTAAAGAAATCTTCAGATCG 273  
QY 121 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGGAAATGTCGGACCTC 180  
Db 274 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGGAAATGTCGGACCTC 333  
QY 181 CTCTGATGATATTTATGAGACATCGCTGGCTGGCTACATTCATTCAGATCTTTGAA 240  
Db 334 CTCTGATGATATTTATGAGACATCGCTGGCTGGCTACATTCATTCAGATCTTTGAA 393  
QY 241 AACATGAACTCTGCAACCTCTCTGGAGAAAGGCAAGGATGACATGAAAA----- 289  
Db 394 AACATGAACTCTGCAACCTCTCTGGAGAAAGGCAAGGATGACATGAAAAGACATTCACA 453  
QY 290 ----- 289  
Db 454 GAAATCTCTGAACGACATGACATGACCAAGGACCAAGCAAGAAAAAGTCCACAGAAATT 513  
QY 290 -----AAATTTCA 297  
Db 514 TCACAGCTGTGCAACAAAGATATGTGCCACAGCTGACAGACAAAAGAACAGAAATTTCA 573  
QY 298 CAAGCTATGAAACAAGAAAGTGGCCACAGACAGACAGACAGAAACAAAGAAATTTCA 357  
Db 574 CAAGCTATGAAACAAGAAAGTGGCCACAGACAGACAGACAGAAACAAAGAAATTTCA 633  
QY 358 GCTATGAAACAAGAAAGTGGCCACAGACAGACAGACAGAAACAAAGGACATGAGAGTAC 417  
Db 634 GCTATGAAACAAGAAAGTGGCCACAGACAGACAGACAGAAACAAAGGACATGAGAGTAC 693  
QY 418 ACATGGGACTACAGAGTCAAGTATGACCAATTCGCTGAGAGAGATGTAAGTCTGCT 477  
Db 694 ACATGGGACTACAGAGTCAAGTATGACCAATTCGCTGAGAGAGATGTAAGTCTGCT 753  
QY 478 AGTTTGAACACCTGCTGCTGACCTGGCCGGAATGCAACGTTGGCTGTGCTTTGAT 537  
Db 754 AGTTTGAACACCTGCTGCTGACCTGGCCGGAATGCAACGTTGGCTGTGCTTTGAT 813  
QY 538 TCAGACGGGTGGGCTTCGGGCTGCAAGGCTGCTGCAAGGAAATGCAAGAAATTGGG 597  
Db 814 TCAGACGGGTGGGCTTCGGGCTGCAAGGCTGCTGCAAGGAAATGCAAGAAATTGGG 873  
QY 598 AAATCGGCTTACGCAAGAGATCGTGTGCTGGGCGCAAGGTGACCTTACCAAGGA 657  
Db 874 AAATCGGCTTACGCAAGAGATCGTGTGCTGGGCGCAAGGTGACCTTACCAAGGA 933  
QY 658 ATGTTCTCTACAGTCTTTCTCTCCGCTTGAAGATGCAAGCGGAAGAGAGACAGT 717  
Db 934 ATGTTCTCTACAGTCTTTCTCTCCGCTTGAAGATGCAAGCGGAAGAGAGACAGT 993  
QY 718 GTCACAGATTCATCTCAGGAGATGCGCAGACTCCAGGCTCCGAGTACGGAATCATG 777  
Db 994 GTCACAGATTCATCTCAGGAGATGCGCAGACTCCAGGCTCCGAGTACGGAATCATG 1053  
QY 778 TCCCGACCAAGAAAGCTGTGTTGATCATTTGATGACGTTTCAGATGACCTGGGCTGTCTC 837  
Db 1054 TCCCGACCAAGAAAGCTGTGTTGATCATTTGATGACGTTTCAGATGACCTGGGCTGTCTC 1113  
QY 838 AACATGACACAAAGCTCTGCAAGAGCTGGGCTGGAAGACAGCTCCGTTACCTCTATA 897  
Db 1114 AACATGACACAAAGCTCTGCAAGAGCTGGGCTGGAAGACAGCTCCGTTACCTCTATA 1173  
QY 898 GCGAGTCTGCTGAGAAAGTCTGCTCCGCTGAGTCTTCCGATGCTCAAGTCAAGAAC 957  
Db 1174 GCGAGTCTGCTGAGAAAGTCTGCTCCGCTGAGTCTTCCGATGCTCAAGTCAAGAAC 1233  
QY 958 GTGGGACAGAGAACTCAAGTCAAGAGGTGCTGCTCCCGTTACCTGTATGAGAGA 1017

Db 1234 GTGGGACAGAGAACTCAAGTCAAGAGTGTGCTCCCGTTACCTGTATGAGAGA 1293  
QY 1018 ATCTCCGGGAAACAABAATCCACTTGTCTTGAAGCGGAGATGCTGACATCAGAG 1077  
Db 1294 ATCTCCGGGAAACAABAATCCACTTGTCTTGAAGCGGAGATGCTGACATCAGAG 1353  
QY 1078 ACACAAGGTTGCTGTCATCATGAAACAAGTGTGCTGACAGTCCAGGTGCC 1137  
Db 1354 ACACAAGGTTGCTGTCATCATGAAACAAGTGTGCTGACAGTCCAGGTGCC 1413  
QY 1138 GCGGTGGGCTCTCATCTGCTGTGCTCCGCTGACAGACGTTGGGGAGAGCGTC 1197  
Db 1414 GCGGTGGGCTCTCATCTGCTGTGCTCCGCTGACAGACGTTGGGGAGAGCGTC 1473  
QY 1198 GCGGCTTCAACCAAAAGCTCACAGGCTGCAAGCGCTTTGTTGTTCAATCAGCTCAC 1257  
Db 1474 GCGGCTTCAACCAAAAGCTCACAGGCTGCAAGCGCTTTGTTGTTCAATCAGCTCAC 1533  
QY 1258 CTTGAGAGCGTGTCCGCGCTGTCTCAATCTGAGAGAAAGATTTCTGAAAGCGCTTC 1317  
Db 1534 CTTGAGAGCGTGTGTCCGCGCTGTCTCAATCTGAGAGAAAGATTTCTGAAAGCGCTTC 1593  
QY 1318 TGCGTATGCTGTGAGAGGAGTGTGGAATGAGAGTCAAGTGTGATGTGACGACCTC 1377  
Db 1594 TGCGTATGCTGTGAGAGGAGTGTGGAATGAGAGTCAAGTGTGATGTGACGACCTC 1653  
QY 1378 ATGCTTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTGTTTCAATGAAACATCTTCTC 1437  
Db 1654 ATGCTTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTGTTTCAATGAAACATCTTCTC 1713  
QY 1438 CCAGACAGCCACTGTGAGGATCTACACCTTCTTCCACCTCAGTCTCAGAACCTTCTGT 1497  
Db 1714 CCAGACAGCCACTGTGAGGATCTACACCTTCTTCCACCTCAGTCTCAGAACCTTCTGT 1773  
QY 1498 GCGGCTTGTATCTACGTTGAGAGGCGCTGGAATGAGCCAGCTCTCTGCTGTGAC 1557  
Db 1774 GCGGCTTGTATCTACGTTGAGAGGCGCTGGAATGAGCCAGCTCTCTGCTGTGAC 1833  
QY 1558 GTTGAAGAAACAAGAGTCCATGAGAGCTTAAACAGGCAAGGCTTCCATTCACCTCCT 1617  
Db 1834 GTTGAAGAAACAAGAGTCCATGAGAGCTTAAACAGGCAAGGCTTCCATTCACCTCCT 1893  
QY 1618 TGATGAAAGCGTTCTTGTGTGCTGTGAGCGAAGCTTAAAGGAGGCACTGGAAGTGC 1677  
Db 1894 TGATGAAAGCGTTCTTGTGTGCTGTGAGCGAAGCTTAAAGGAGGCACTGGAAGTGC 1953  
QY 1678 CTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGCTTGTGACCTGGGCTCTCTG 1737  
Db 1954 CTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGCAAGCTTGTGACCTGGGCTCTCTG 2013  
QY 1738 TTTGGGTGAGAGCCTTATGCAACCAAGGAGAGACCCCTGAGAGGCTTCCACTGTCTT 1797  
Db 2014 TTTGGGTGAGAGCCTTATGCAACCAAGGAGAGACCCCTGAGAGGCTTCCACTGTCTT 2073  
QY 1798 TTTGAGACTCAAGCAAGAGTTGTGCTGTGCTTAAACAGCTTCAAGAGTGTGG 1857  
Db 2074 TTTGAGACTCAAGCAAGAGTTGTGCTGTGCTTAAACAGCTTCAAGAGTGTGG 2133  
QY 1858 CTTTCGATTAACAGAACTGTGACTTATGATCTTCTTGTGCTTCAGACCTGTCCG 1917  
Db 2134 CTTTCGATTAACAGAACTGTGACTTATGATCTTCTTGTGCTTCAGACCTGTCCG 2193  
QY 1918 TATTTGGGAAAAATTCGGGTGTGATGCAAAAGGATCTTCCCAAGATGATGCTCCGCTGAG 1977  
Db 2194 TATTTGGGAAAAATTCGGGTGTGATGCAAAAGGATCTTCCCAAGATGATGCTCCGCTGAG 2253  
QY 1978 GCATGCTGTGTGCTCTTATGATGAGTGGGGAATAAGCCCTCATTTGAGAGCAGTGGGAA 2037  
Db 2254 GCATGCTGTGTGCTCTTATGATGAGTGGGGAATAAGCCCTCATTTGAGAGCAGTGGGAA 2313  
QY 2038 GATTTCTGCTTCATGCTTTGGACCCACCAAGCTGTGGGAGCTGTGACCTGGGAGAGC 2097

Db 2314 GATTTCCTCCATGCTGGACCCACCACTGCGGAGCTGAGACTTGGGACAGC 2373  
Qy 2098 ATCCGACAGACGGGCGCCATGAAAGACCTGTGTGCGCACTGAGGATCTCCACCTGCAAG 2157  
Db 2374 ATCCGACAGACGGGCGCCATGAAAGACCTGTGTGCGCACTGAGGATCTCCACCTGCAAG 2433  
Qy 2158 ATACAGACCTGATGTTAGAAATGCAAGATTAACCTGTGTGCGCACTGAGGATCTCCACCTGCAAG 2217  
Db 2434 ATACAGACCTGATGTTAGAAATGCAAGATTAACCTGTGTGCGCACTGAGGATCTCCACCTGCAAG 2493  
Qy 2218 ATGCTGATGCGCAACCGTAACCTGATGCTTCAACTTGGAGGACCCACCTGAAAGAA 2277  
Db 2494 ATGCTGATGCGCAACCGTAACCTGATGCTTCAACTTGGAGGACCCACCTGAAAGAA 2553  
Qy 2278 GAGGATGTAAGATGCGGCTGTGAAAGCTTAAACACCCCAATGTTTGTGAGTCTTTG 2337  
Db 2554 GAGGATGTAAGATGCGGCTGTGAAAGCTTAAACACCCCAATGTTTGTGAGTCTTTG 2613  
Qy 2338 AGGCTGATGCTGTGATTTGACCCATGCTGTAAAGATCTCCCAATCTTTAG 2397  
Db 2614 AGGCTGATGCTGTGATTTGACCCATGCTGTAAAGATCTCCCAATCTTTAG 2673  
Qy 2398 ACCCTCCCGAGCTGAAATCTGTGAGCTGCGAGAAACAGGTGACAGACAGGAGTA 2457  
Db 2674 ACCCTCCCGAGCTGAAATCTGTGAGCTGCGAGAAACAGGTGACAGACAGGAGTA 2733  
Qy 2458 ATGCTCTGATGATGCTGTGAGGATCTCCAGTGGCGCTGCAAGAGCTGATTTAGAG 2517  
Db 2734 AGGCTCTGATGATGCTGTGAGGATCTCCAGTGGCGCTGCAAGAGCTGATTTAGAG 2793  
Qy 2518 GACTGAGCATACAGCCAGCGGTTGCGAGAGCTGCGCTGCAAGCTGCTGCAAGACCG 2577  
Db 2794 GACTGAGCATACAGCCAGCGGTTGCGAGAGCTGCGCTGCAAGCTGCTGCAAGACCG 2853  
Qy 2578 AGCTTGAACACCTGTGCTTATCAACACAGCTGCGGAGCAAGAGTGAATCTACTG 2637  
Db 2854 AGCTTGAACACCTGTGCTTATCAACACAGCTGCGGAGCAAGAGTGAATCTACTG 2913  
Qy 2638 TGTGATCATGAGGCTTCCCACTGTGATGCTGCAAGAGCTGATGCTGATGCTGAC 2697  
Db 2914 TGTGATCATGAGGCTTCCCACTGTGATGCTGCAAGAGCTGATGCTGATGCTGAC 2973  
Qy 2698 CTGGAACAGGCTGCGGCTGCTTCTTGTGCACTTGGCTTAAAGGTTAACTGAGCTGAG 2757  
Db 2974 CTGGAACAGGCTGCGGCTGCTTCTTGTGCACTTGGCTTAAAGGTTAACTGAGCTGAG 3033  
Qy 2758 CACTGAGCTTGAATGAAACCTGTGAAAGCAATGCGGTGAAGCTTGTGCGAGCTG 2817  
Db 3034 CACTGAGCTTGAATGAAACCTGTGAAAGCAATGCGGTGAAGCTTGTGCGAGCTG 3093  
Qy 2818 ATGAGAGAACATCTTGTGATCTCAAGACCTGAGATTGTAAGTGTGATCTGACCGC 2877  
Db 3094 ATGAGAGAACATCTTGTGATCTCAAGACCTGAGATTGTAAGTGTGATCTGACCGC 3153  
Qy 2878 GGTGCTGTGAGAGCTGCTGCTGTGATCTGAGAGAGAGCACTTGAAGGCTGTGAT 2937  
Db 3154 GGTGCTGTGAGAGCTGCTGCTGTGATCTGAGAGAGAGCACTTGAAGGCTGTGAT 3213  
Qy 2938 CTCAAGGACAAATGCTGCGGTGAGCGGTGCTGCGCTGTGCGAGGAGCTGAAGCA 2997  
Db 3214 CTCAAGGACAAATGCTGCGGTGAGCGGTGCTGCGCTGTGCGAGGAGCTGAAGCA 3273  
Qy 2998 AAGAACAGTGTCTGACAGACCTGCGGTTGAAGGAGTGTGACCTGATTTGCTGT 3057  
Db 3274 AAGAACAGTGTCTGACAGACCTGCGGTTGAAGGAGTGTGACCTGATTTGCTGT 3333  
Qy 3058 GAGGACATCTCTTGGGCTTCTGCGCAACCGGCACTTGAACAGTCTTAACTGTGTGAG 3117  
Db 3334 GAGGACATCTCTTGGGCTTCTGCGCAACCGGCACTTGAACAGTCTTAACTGTGTGAG 3393  
Qy 3118 AATACTTCACTGCGCAAGAGATGATGAGCTGTGTGCGCTTGTGCTGCTGCTGCT 3177  
Db 3394 AATACTTCACTGCGCAAGAGATGATGAGCTGTGTGCGCTTGTGCTGCTGCTGCT 3453

Qy 3178 AACTTACAGATTAATTGGGCTGTGGAATGCGAGTAACCTGTGCAATTAAGAGCTGCG 3237  
Db 3454 AACTTACAGATTAATTGGGCTGTGGAATGCGAGTAACCTGTGCAATTAAGAGCTGCG 3513  
Qy 3238 GAGGAAGTCAGCTACTCAAGCCCGAGTGTGAAATGAGCGTGTGATCTTTGAT 3297  
Db 3514 GAGGAAGTCAGCTACTCAAGCCCGAGTGTGAAATGAGCGTGTGATCTTTGAT 3573  
Qy 3298 GAAATGACCGAC 3312  
Db 3574 GAAATGACCGGTAC 3588  
RESULT 5  
US-10-677-943-23  
; Sequence 23, Application US/10677943  
; Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zhi-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT APPLICATION NUMBER: US/10/677,943  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US02/09776  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 23  
; LENGTH: 3900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3600)  
US-10-677-943-23  
Query Match 78.5%; Score 3166.2; DB 18; Length 3900;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;  
Qy 1 ATGAGAGAGCAAAATGCTCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 60  
Db 154 ATGAGAGAGCAAAATGCTCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 213  
Qy 61 CTAGCAAGGAAGATTTTCAAGCAATTCAGGAATTTACTTAAGGAATCTTCAAGATCG 120  
Db 214 CTAGCAAGGAAGATTTTCAAGCAATTCAGGAATTTACTTAAGGAATCTTCAAGATCG 273  
Qy 121 ACCACATGCTTATTTCCACAGTTTGAATCGAGATGCAAGTGAATGTCTGCACTC 180  
Db 274 ACCACATGCTTATTTCCACAGTTTGAATCGAGATGCAAGTGAATGTCTGCACTC 333  
Qy 181 CTCTTCAATGATTAATGAGCATGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Db 334 CTCTTCAATGATTAATGAGCATGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 393  
Qy 241 AACATGAACCTGCGAACCTCTGCGAGAGGACAGGAGTGAAGTGAAGTGAAGTGAAG 289  
Db 394 AACATGAACCTGCGAACCTCTGCGAGAGGACAGGAGTGAAGTGAAGTGAAGTGAAG 453  
Qy 290 ----- 289  
Db 454 GAAATCTGTAAGCAAGATGATGACCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGT 513



QY 290 -----AAATTCa 297  
Db 514 TCACAAGCTGTGCAACAAGATAGTCCACAGCTGCAGAGACAAAAGAAACAGAAATTTTCA 573  
QY 298 CAAGCTATGGAACAAGAGGTGCCACAGACAGACAGACAAAGAAACAAGAAATTTTCAAA 357  
Db 574 CAAGCTATGGAACAAGAGGTGCCACAGACAGACAGACAAAGAAACAAGAAATTTTCAAA 633  
QY 358 GCTATNGGAACAAGAGGTGCCACAGACAGACAGACAAAGAAACAAGAAACAAGAGGTAC 417  
Db 634 GCTATNGGAACAAGAGGTGCCACAGACAGACAGACAAAGAAACAAGAAACAAGAGGTAC 693  
QY 418 ACATGGAGCTACAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGAGATAGTACGTGT 477  
Db 694 ACATGGAGCTACAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGAGATAGTACGTGT 753  
QY 478 AGTTTGAACACATGCTGCTGACTGCGCCGGAATGCAACGTTGGCTGTGTCTTTGAT 537  
Db 754 AGTTTGAACACATGCTGCTGACTGCGCCGGAATGCAACGTTGGCTGTGTCTTTGAT 813  
QY 538 TCAGACCGGTGGGGCTTCGGCTTCGCACGGTGTTCTGCAACGGAATGCAAGAAATTTGG 597  
Db 814 TCAGACCGGTGGGGCTTCGGCTTCGCAACGGTGTTCTGCAACGGAATGCAAGAAATTTGG 873  
QY 598 AAATCGGCTTAGCCAGAAAGATCGTGTGCTGTGGCGCAAGGTGACTACCAAGGGA 657  
Db 874 AAATCGGCTTAGCCAGAAAGATCGTGTGCTGTGGCGCAAGGTGACTACCAAGGGA 933  
QY 658 ATGTTCTCTAAGTCTTCTTCTCCCGTTAGAGATGACGCGGAAGAAGAGACAGT 717  
Db 934 ATGTTCTCTAAGTCTTCTTCTCCCGTTAGAGATGACGCGGAAGAAGAGACAGT 993  
QY 718 GTCCAGAGTTCACTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTACGGAATGCAATG 777  
Db 994 GTCCAGAGTTCACTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTACGGAATGCAATG 1053  
QY 778 TCCCGACCAAGAAAGCTGTTGTTCAATTCAGCGTTTCAGTACCTGGGCTGTCTC 837  
Db 1054 TCCCGACCAAGAAAGCTGTTGTTCAATTCAGCGTTTCAGTACCTGGGCTGTCTC 1113  
QY 838 AACCAATGACAAAGCTCTGCAAGAAGCTGGGCTGGAAGACAGCTCCGTTACCTCTATA 897  
Db 1114 AACCAATGACAAAGCTCTGCAAGAAGCTGGGCTGGAAGACAGCTCCGTTACCTCTATA 1173  
QY 898 CGCAGCTGCTGAGGAAGTCTGCTCCCTGAGTCCCTGATGTCCTGATGTCACCTGCAAG 957  
Db 1174 CGCAGCTGCTGAGGAAGTCTGCTCCCTGAGTCCCTGATGTCCTGATGTCACCTGCAAG 1233  
QY 958 GTGGCACAAGAAAGCTCAAGTCAAGAGTCTGCTCCCGTTACCTGTTAGTTAGAGA 1017  
Db 1234 GTGGCACAAGAAAGCTCAAGTCAAGAGTCTGCTCCCGTTACCTGTTAGTTAGAGA 1293  
QY 1018 ATCTCCGGGGAACAAGAAATTCATTGCTCTTGAGCGCGGGATTTGTGACATCAAG 1077  
Db 1294 ATCTCCGGGGAACAAGAAATTCATTGCTCTTGAGCGCGGGATTTGTGAGATCAAG 1353  
QY 1078 ACACAAGGTTGCGGCAATCAATGAACAACGTTGAGTCTGCAACAGTGCAGGTTGCC 1137  
Db 1354 ACACAAGGTTGCGGCAATCAATGAACAACGTTGAGTCTGCAACAGTGCAGGTTGCC 1413  
QY 1138 GCCGTGGGCTCTCATCTGCTGCGCCCTGACAGCTGACAGACGTGTGGGGAGAGCGTC 1197  
Db 1414 GCCGTGGGCTCTCATCTGCGTGGCCCTGACAGCTGACAGAGTGTGTGGGGAGAGCGTC 1473  
QY 1198 GCCCCTTCAACAAACGCTCAACAGGCTGACGCGCTTTTGTGTTTCAATGACTCAC 1257  
Db 1474 GCCCCTTCAACAAACGCTCAACAGGCTGACGCGCTTTTGTGTTTCAATGACTCAC 1533  
QY 1258 CCTGAGGCGTGTCCGCGCTGTCTCAATCTGAGGAAAGAGTGTCTGAAAGGCTTC 1317  
Db 1534 CCTGAGGCGTGTCCGCGCTGTCTCAATCTGAGGAAAGAGTGTCTGAAAGGCTTC 1593  
QY 1318 TGCCGTATGTGTGAGGAGTGTGAATAGGAAGTCAAGTGTGATGTGACGACCTC 1377

Db 1594 TGCCGTATGTGTGAGGAGTGTGAATAGGAATGCAAGTGTGATGTGAGACCTC 1653  
QY 1378 ATGTTCAAGACCTCGGGAGTCTGAGACTCCGTGCTGTGTTCAATGAACATCTTCTC 1437  
Db 1654 ATGTTCAAGACCTCGGGAGTCTGAGACTCCGTGCTGTGTTCAATGAACATCTTCTC 1713  
QY 1438 CCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAAGTCTTCAGAACTTCTGT 1497  
Db 1714 CCAGACAGCCACTGTGAGGAGTACTACACTTCTTCCACCTCAAGTCTTCAGAACTTCTGT 1773  
QY 1498 GCCGCTTTGATCAAGTGTAGAGGGCTGGAATCGAAGCCAGCTCTGCTGTGAC 1557  
Db 1774 GCCGCTTTGATCAAGTGTAGAGGGCTGGAATCGAAGCCAGCTCTGCTGTGAC 1833  
QY 1558 GTTGAAGAACAAGAGTCAATGAGACTTAAACAGCAGGCTTCCATATCCACTCCCTT 1617  
Db 1834 GTTGAAGAACAAGAGTCAATGAGACTTAAACAGCAGGCTTCCATATCCACTCCCTT 1893  
QY 1618 TGAATGAAGCGTTCTTGTGGCTGTGAGCGGAACGTAAGAGAGCCACTGAGGTC 1677  
Db 1894 TGAATGAAGCGTTCTTGTGTGGCTGTGAGCGGAACGTAAGAGAGCCACTGAGGTC 1953  
QY 1678 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAAAGCTTGTGACTGCGTCTGTG 1737  
Db 1954 CTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAAAGCTTGTGACTGCGTCTGTG 2013  
QY 1738 TTGGGTCAAGCGCTTAATGCCACACCCCAAGAACACCTCGACGCTTCCACTGTCTT 1797  
Db 2014 TTGGGTCAAGCGCTTAATGCCACACCCCAAGAACACCTCGACGCTTCCACTGTCTT 2073  
QY 1798 TTGAGACTCAAGACAAGAGTTGTTGCTTGGCATTTAAACAGCTTCAAGAGTGTG 1857  
Db 2074 TTGAGACTCAAGACAAGAGTTGTTGCTTGGCATTTAAACAGCTTCAAGAGTGTG 2133  
QY 1858 CTTCGATTAACAAGAACTTGACATCTTCTGTGCTTCAGACCTGTCCG 1917  
Db 2134 CTTCGATTAACAAGAACTTGACATCTTCTGTGCTTCAGACCTGTCCG 2193  
QY 1918 TATTTGGGGAATAATTCGGGTGATGTCAAGAGATCTTCCCAAGATGAGTCCGCTGAG 1977  
Db 2194 TATTTGGGGAATAATTCGGGTGATGTCAAGAGATCTTCCCAAGATGAGTCCGCTGAG 2253  
QY 1978 GCATGCTCGTGTGCTCTCTATGATGAGCGGGAATAAGACCTCATTTGAGAGACGTGGAA 2037  
Db 2254 GCATGCTCGTGTGCTCTCTATGATGAGCGGGAATAAGACCTCATTTGAGAGACGTGGAA 2133  
QY 2038 GATTTCTGCTCATGCTTTGGCACCCACCAACCTGCGGACGTGACCTTGGGACAGAC 2097  
Db 2314 GATTTCTGCTCATGCTTTGGCACCCACCAACCTGCGGACGTGACCTTGGGACAGAC 2373  
QY 2098 ATCTGACAGCGGGCCATGAAACCTGTGTGTGCAAGCTGAGGATCCCACTTGCAG 2157  
Db 2374 ATCTGACAGCGGGCCATGAAACCTGTGTGTGCAAGCTGAGGATCCCACTTGCAG 2433  
QY 2158 ATACAGACCCGATGTTTGAATAATGACAAATTACCCCTGGTGTGACAGACCTTGGAGA 2217  
Db 2434 ATACAGACCCGATGTTTGAATAATGACAAATTACCCCTGGTGTGACAGACCTTGGAGA 2493  
QY 2218 ATGTCATGGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCCACTGAAGGA 2277  
Db 2494 ATGTCATGGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCCACTGAAGGA 2553  
QY 2278 GAGGATGTAAGAGTGGGTGTGAAGCTTAAACACCCCAAAATTTGTTGGAATCTTTG 2337  
Db 2554 GAGGATGTAAGAGTGGGTGTGAAGCTTAAACACCCCAAAATTTGTTGGAATCTTTG 2613  
QY 2338 AGGCTGAATGTCTGTGATTAACCATGCGCTTAACCTGAAGATCTCCAAATCTTACG 2397  
Db 2614 AGGCTGAATGTCTGTGATTAACCATGCGCTTAACCTGAAGATCTCCAAATCTTACG 2673  
QY 2398 ACTTCCCAAGCTGAATCTGTGAGCTGTGACAGAAACAAGTGAAGACAGGAGTA 2457



```

Db      2674 ACCGCCCGAGCTGAAATCTCTGAGCCTGGCAGAAACAAGGTGACAGACCAGGAGTA 2733
Qy      2458 ATGGCTCTAGTGAAGCCTTGAAGTCTCCAGTGGCCCTGGAAGAGTGTACTGTGAG 2517
Db      2734 AGGCTCTAGTGAAGCCTTGAAGGCTCTCCAGTGGCCCTGGAAGAGTGTACTGTGAG 2793
Qy      2518 GACTGTGATCAACAGCCAGGCTTGCAGAGTCTGGCCTCAAGCCCTCGTCAGAACCGG 2577
Db      2794 GACTGTGATCAACAGCCAGGCTTGCAGAGTCTGGCCTCAAGCCCTCGTCAGAACCGG 2853
Qy      2578 AGCTTGACACACCTGTCTATCCAAACAAGCCTGGGGAAGAGGTGTAAATCTACTG 2637
Db      2854 AGCTTGACACACCTGTCTATCCAAACAAGCCTGGGGAAGAGGTGTAAATCTACTG 2913
Qy      2638 TGTTCATCATGAGGTTTCCCACTGTGTCTGCAAGAGCTGATGTGTAATCACTGCCAC 2697
Db      2914 TGTTCATCATGAGGTTTCCCACTGTGTCTGCAAGAGCTGATGTGTAATCACTGCCAC 2973
Qy      2698 CTGGAACAGGCTGGCTGTGTTTCTTGTGACCTTGTGAGCTTATGGGTAACTATGGCTGAG 2757
Db      2974 CTGGAACAGGCTGGCTGTGTTTCTTGTGACCTTGTGAGCTTATGGGTAACTATGGCTGAG 3033
Qy      2758 CACTGAGCCTTAGACATGAACCTGTGAAACAATGCGGTGAAGCTTCTGTGCGAGGTC 2817
Db      3034 CACTGAGCCTTAGACATGAACCTGTGAAACAATGCGGTGAAGCTTCTGTGCGAGGTC 3093
Qy      2818 ATGAAGAACAATCTTGTGTATCTCCAGGACCTGGAAGTGTAAAGTGTATCTCACCGCC 2877
Db      3094 ATGAAGAACAATCTTGTGTATCTCCAGGACCTGGAAGTGTAAAGTGTATCTCACCGCC 3153
Qy      2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTGAGGAGACGACCTGGAAGAGCCTGAT 2937
Db      3154 GCGTCTGTGAGAGTCTGTCTGTGTGATCTGAGGAGACGACCTGGAAGAGCCTGAT 3213
Qy      2938 CTCAAGCAATAGCCTGTGGTGAACGATGGGCTTGTGCTGCTGTGCGAGGGACTGAAGCA 2997
Db      3214 CTCAAGCAATAGCCTGTGGTGAACGATGGGCTTGTGCTGCTGTGCGAGGGACTGAAGCA 3273
Qy      2998 AAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCAATGTGACTGACTTGTGATGTGCTGT 3057
Db      3274 AAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCAATGTGACTGACTTGTGATGTGCTGT 3333
Qy      3058 GAGGACTCTCTTGGCCCTTCTGCAACCGGCAATCTGACCAAGCTTAAACCTGGTGCAG 3117
Db      3334 GAGGACTCTCTTGGCCCTTCTGCAACCGGCAATCTGACCAAGCTTAAACCTGGTGCAG 3393
Qy      3118 AATACTTCACTGCTCCAAAGAAATGAAGTGTGTGGCCTTGTGCTGTCCACAGTCT 3177
Db      3394 AATACTTCACTGCTCCAAAGAAATGAAGTGTGTGGCCTTGTGCTGTCCACAGTCT 3453
Qy      3178 AACTTACAGATTAATGGCTGTGGAATGGCAATGACCTGTGCAATAAAGAAAGCTGTG 3237
Db      3454 AACTTACAGATTAATGGCTGTGGAATGGCAATGACCTGTGCAATAAAGAAAGCTGTG 3513
Qy      3238 GAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTTGACGTAATTTGATTTTGTAT 3297
Db      3514 GAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTTGACGTAATTTGATTTTGTAT 3573
Qy      3298 GAAAGTGAACCGACAC 3312
Db      3574 GAAAGTGAACCGGTAC 3588

```

RESULT 6  
US-10-216-645-3

; Sequence 3, Application US/10216645  
; Publication No. US20030125282A1

; GENERAL INFORMATION:

; APPLICANT: WEISS, BERTRAM

; APPLICANT: LESSL, MONIKA

; APPLICANT: PETERS-KOTTIG, MICHAEL

; APPLICANT: BECKMANN, GEORG

; TITLE OF INVENTION: HUMAN MATER PROTEINS

```

; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3830
; ORGANISM: Homo sapiens
; TYPE: DNA
US-10-216-645-3

Query Match      77.8%; Score 3139; DB 15; Length 3830;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 3282; Conservative 0; Mismatches 30; Indels 105; Gaps 2;

Qy      1 ATGAAGAAGACAAATTCCTCACTTTTCAAGCTACGGGCTGCAATGTGTCTATGAG 60
Db      1 ATGAAGAAGACAAATTCCTCACTTTTCAAGCTACGGGCTGCAATGTGTCTATGAG 60
Qy      61 CTAGAACAAGAAATTTCAACATTCAGAAATTAATAAGAAATTTCTCAGAAATCG 120
Db      61 CTAGAACAAGAAATTTCAACATTCAGAAATTAATAAGAAATTTCTCAGAAATCG 120
Qy      121 ACCAATGCTCTATTTCCACAGTTTGAATTCAGAAATGCCAAGTGTATGTTCGACATC 180
Db      121 ACCAATGCTCTATTTCCACAGTTTGAATTCAGAAATGCCAAGTGTATGTTCGACATC 180
Qy      181 CTCTGCAATGATTAATTAAGAGCATCGCTGGCTGGGCTACGTCATAGCATCTTTGA 240
Db      181 CTCTGCAATGATTAATTAAGAGCATCGCTGGCTGGGCTACGTCATAGCATCTTTGA 240
Qy      241 AACATGAACCTGCGAACCTCTCGGAGAAAGCACGGATGACATGAAAAAATTTACAA 300
Db      241 AACATGAACCTGCGAACCTCTCGGAGAAAGCACGGATGACATGAAAAAATTTACAA 300
Qy      301 GCT-----ATGACACAGAGGTGCAACAGCAGCAGACAGAAAGCA- 344
Db      301 GATCTGAAGAACGATGACATGACCAAGAACCAAGAAAGATGTCAGAAAAATTA 360
Qy      345 ----- 344
Db      361 TATGCAATGACTAAGCTTATCTTGGGGGTGTGACATCTCTGACTCGAATTAATTAAC 420
Qy      345 -----AGAAATTTCAACAGCTATGAACCAAGAGT 375
Db      421 AAGTATGTGAATTCATTTCTTTTGCAAGAAATTTCAACAGCTATGAACCAAGAGT 480
Qy      376 GCCACAGCAGACAGACAGAAAGAACAGACATGAGGTGACACATGGGACTTACAAAGT 435
Db      481 GCCACAGCAGACAGACAGAAAGAACAGACATGAGGTGACACATGGGACTTACAAAGT 540
Qy      436 CACGTGATGACCAATTTGCTGAGAGAGAGATGACGTCTAGTTTGAATAAACAATCT 495
Db      541 CACGTGATGACCAATTTGCTGAGAGAGAGATGACGTCTAGTTTGAATAAACAATCT 600
Qy      496 GCTGATGGCCGGAATCAACAGTGTGCTGTGTTTGAATTCAGACCGGTGGGGCTTC 555
Db      601 GCTGATGGCCGGAATCAACAGTGTGCTGTGTTTGAATTCAGACCGGTGGGGCTTC 660
Qy      556 CGGCTCGACAGGTGTCTGCAAGGAAAGTCAAGAAATTTGGAATTCGCTTACCCAGA 615
Db      661 CGGCTCGACAGGTGTCTGCAAGGAAAGTCAAGAAATTTGGAATTCGCTTACCCAGA 720
Qy      616 AGGATCGTGTGTGCTGGGGCAAGGTGACTTACAGAGGAATTTCTCTAGTCTTC 675
Db      721 AGGATCGTGTGTGCTGGGGCAAGGTGACTTACAGAGGAATTTCTCTAGTCTTC 780
Qy      676 TTCTCTCCCGTTAAGAGATGACAGGAAAGAGAGACAGTGTACAGAGTTCAATCTCC 735
Db      781 TTCTCTCCCGTTAAGAGATGACAGGAAAGAGAGACAGTGTCAAGAGTTCAATCTCC 840

```

Qy 736 AGGAGTGGCAGACTCCCAAGGCTCCGGTGAAGGATCATGTCCCGACAGAAAAGGCTG 795  
Db 841 AGGAGTGGCAGACTCCCAAGGCTCCGGTGAAGGATCATGTCCCGACAGAAAAGGCTG 900  
Qy 796 TTGTTCATCATTTGACGGTTTGTGATGACTGGGCTCTGTCTTCACATGACACAAAGCTC 855  
Db 901 TTGTTCATCATTTGACGGTTTGTGATGACTGGGCTCTGTCTTCACATGACACAAAGCTC 960  
Qy 856 TGCAGAGCTGGGGCTGAGAGAGAGCTCGTTCAACCTTCATACGAGCTGTGAGAG 915  
Db 961 TGCAGAGCTGGGGCTGAGAGAGAGCTCGTTCAACCTTCATACGAGCTGTGAGAG 1020  
Qy 916 GTCCGTCTCCCTGAGTCCCTCTGATCGTCAACCGTCAGAGAGCTGGGACAGAGAGCTC 975  
Db 1021 GTCCGTCTCCCTGAGTCCCTCTGATCGTCAACCGTCAGAGAGCTGGGACAGAGAGCTC 1080  
Qy 976 AAGTCAGAGTGTGTCTTCCCGTTACCTGTTAGTAGAGANCTCCGGGAAACAAAGA 1035  
Db 1081 AAGTCAGAGTGTGTCTTCCCGTTACCTGTTAGTAGAGANCTCCGGGAAACAAAGA 1140  
Qy 1036 ATCCACTGTCTTGAAGCGGGGATGTGTGAGCATGAGAGACAGAGGGTGTGTCG 1095  
Db 1141 ATCCACTGTCTTGAAGCGGGGATGTGTGAGCATGAGAGACAGAGGGTGTGTCG 1200  
Qy 1096 ATCATGAACAAACCGTGAAGTGTCTGACCAAGTGCAGAGTGCAGCGGCTCTCTCATC 1155  
Db 1201 ATCATGAACAAACCGTGAAGTGTCTGACCAAGTGCAGAGTGCAGCGGCTCTCTCATC 1260  
Qy 1156 TGCCTGCTCTGACACTGACAGACGTGTGGGGAGAGCGTGCCTCTTACCAACAAAG 1215  
Db 1261 TGCCTGCTCTGACACTGACAGACGTGTGGGGAGAGCGTGCCTCTTACCAACAAAG 1320  
Qy 1216 CTCACAGGCTCGACAGCGCTTTGTGTGTTTCAATGAGCTCAACCCCTCGAGGGTGTGTCG 1275  
Db 1321 CTCACAGGCTCGACAGCGCTTTGTGTGTTTCAATGAGCTCAACCCCTCGAGGGTGTGTCG 1380  
Qy 1276 CGCTGTCTCAATCTGAGAGAAAGAGTTGTCTGAGAGCGCTTCTGCGGTGTGTGAG 1335  
Db 1381 CGCTGTCTCAATCTGAGAGAAAGAGTTGTCTGAGAGCGCTTCTGCGGTGTGTGAG 1440  
Qy 1336 GGAAGTGTGAATAGAGAGTCAAGTGTGTGATGTGTGACGACTTCATGTGTTCAAGAGCTCGGG 1395  
Db 1441 GGAAGTGTGAATAGAGAGTCAAGTGTGTGATGTGTGACGACTTCATGTGTTCAAGAGCTCGGG 1500  
Qy 1396 GAGTGTGAGCTCCGCTCTGTGTTCAATGAACATCTTCTCCAGACAGGACCTGTGAG 1455  
Db 1501 GAGTGTGAGCTCCGCTCTGTGTTCAATGAACATCTTCTCCAGACAGGACCTGTGAG 1560  
Qy 1456 GAGTACTACACCTTCTTCCACTCAGTCTCCAGACTTCTGTGCGGCTTGTACTAGCTG 1515  
Db 1561 GAGTACTACACCTTCTTCCACTCAGTCTCCAGACTTCTGTGCGGCTTGTACTAGCTG 1620  
Qy 1516 TTAGAGGGCTTGAAGATGAGGACAGCTCTGTGCTGTGTACGTTGAGAGACAAAGAGG 1575  
Db 1621 TTAGAGGGCTTGAAGATGAGGACAGCTCTGTGCTGTGTACGTTGAGAGACAAAGAGG 1680  
Qy 1576 TCCATGAGACTTAAACAAGGAGGCTTCCATATCCACTGGCTTGTGAGTAGAGGTTCTTG 1635  
Db 1681 TCCATGAGACTTAAACAAGGAGGCTTCCATATCCACTGGCTTGTGAGTAGAGGTTCTTG 1740  
Qy 1636 TTTGGCTCTGTGAGCGAGAGAGCTGTAAGAGGCCACTGAGAGTCTGTGGGCTGTCCGCTT 1695  
Db 1741 TTTGGCTCTGTGAGCGAGAGAGCTGTAAGAGGCCACTGAGAGTCTGTGGGCTGTCCGCTT 1800  
Qy 1696 CCCCTGGGGTGAACAGAGAGCTTCTGCACTGGGTCTCTGTGTGGGTCAAGAGCTTAAT 1755  
Db 1801 CCCCTGGGGTGAACAGAGAGCTTCTGCACTGGGTCTCTGTGTGGGTCAAGAGCTTAAT 1860  
Qy 1756 GGCACCAACCCAGAGAGACCTGTGACGCTTCCACTGTCTTTTGTGAGACTCAAGAGCAA 1815  
Db 1861 GGCACCAACCCAGAGAGACCTGTGACGCTTCCACTGTCTTTTGTGAGACTCAAGAGCAA 1920  
Qy 1816 GAGTTTGTTCGCTTGGCATTAACAGCTTCCAAAGAGTGTGGCTTCCGATTAACAGAAC 1875

Db 1921 GAGTTTGTTCGCTTGGCATTAACAGCTTCCAAAGAGTGTGGCTTCCGATTAACAGAAC 1980  
Qy 1876 CTGAGCTTGAATAGATCTTCTTGTGCTCTCAGACACTGTCCGTAATTTGGGAAAATTCGG 1935  
Db 1981 CTGAGCTTGAATAGATCTTCTTGTGCTCTCAGACACTGTCCGTAATTTGGGAAAATTCGG 2040  
Qy 1936 GTGATGTCAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGATGTCTGTGGTCCCT 1995  
Db 2041 GTGATGTCAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGATGTCTGTGGTCCCT 2100  
Qy 1996 CTATGATGCGGGATTAAGACCTCTGATTGAGAGAGAGTGGAGATTTCTGTCCATGCTT 2055  
Db 2101 CTATGATGCGGGATTAAGACCTCTGATTGAGAGAGAGTGGAGATTTCTGTCCATGCTT 2160  
Qy 2056 GGCACCAACCCAGACCTGTGCGGAGCTGTGACTGGGACAGACATCTCTGACAGAGCGGGCC 2115  
Db 2161 GGCACCAACCCAGACCTGTGCGGAGCTGTGACTGGGACAGACATCTCTGACAGAGCGGGCC 2220  
Qy 2116 ATGAAGACCCGTGTGCGAAGCTGAGGCAATCCCACTGCAAGATPACAGACCTGATGTTT 2175  
Db 2221 ATGAAGACCCGTGTGCGAAGCTGAGGCAATCCCACTGCAAGATPACAGACCTGATGTTT 2280  
Qy 2176 AGAATGACAGATTACCCCTGTGTGACAGACCTCTGAGAAATCGTCATGGCCAAACGT 2235  
Db 2281 AGAATGACAGATTACCCCTGTGTGACAGACCTCTGAGAAATCGTCATGGCCAAACGT 2340  
Qy 2236 AACCTAAGATCCCTCAACTTGGAGAGCACCACTGAAAGAAAGATGTAAAGATGGCG 2295  
Db 2341 AACCTAAGATCCCTCAACTTGGAGAGCACCACTGAAAGAAAGATGTAAAGATGGCG 2400  
Qy 2296 TGTGAAGCTTAAACACCCAAATGTTTGTGAGAGTCTTGAAGGCTGATGTTGCTGTGGA 2355  
Db 2401 TGTGAAGCTTAAACACCCAAATGTTTGTGAGAGTCTTGAAGGCTGATGTTGCTGTGGA 2460  
Qy 2356 TTGAACCAATGCTTGAACCTGAGATCTCCAAATCTTGAAGCTTCCCAAGCTGAAA 2415  
Db 2461 TTGAACCAATGCTTGAACCTGAGATCTCCAAATCTTGAAGCTTCCCAAGCTGAAA 2520  
Qy 2416 TCTCTGAGCTGTGAGAGAAACAGAGTGTACAGACCAAGGAGTAATGCTCTTCAAGTATGCC 2475  
Db 2521 TCTCTGAGCTGTGAGAGAAACAGAGTGTACAGACCAAGGAGTAATGCTCTTCAAGTATGCC 2580  
Qy 2476 TTGAGAGTCTCCAGTGTGCGCTCTGACAGAGCTGATATCTGAGAGCTGTGGCATCACAGCC 2535  
Db 2581 TTGAGAGTCTCCAGTGTGCGCTCTGACAGAGCTGATATCTGAGAGCTGTGGCATCACAGCC 2640  
Qy 2536 ACGGTTGCGAGAGTGTGAGCTTCAAGCCTGTGTACAGAACCGAGACTTGAACAACCTGTGC 2595  
Db 2641 ACGGTTGCGAGAGTGTGAGCTTCAAGCCTGTGTACAGAACCGAGACTTGAACAACCTGTGC 2700  
Qy 2596 CTATCCAAACAAGCTGTGGGAAACGAAGGTGTAAATCTATCTGTGTGATCCATAGAGCTT 2655  
Db 2701 CTATCCAAACAAGCTGTGGGAAACGAAGGTGTAAATCTATCTGTGTGATCCATAGAGCTT 2760  
Qy 2656 CCCCACTGTATGTGCAAGGCTGATGTGAATGAGGCAACCTGGAACAAGGCTGGGCTGT 2715  
Db 2761 CCCCACTGTATGTGCAAGGCTGATGTGAATGAGGCAACCTGGAACAAGGCTGGGCTGT 2820  
Qy 2716 GGTTTTCTTGAACCTTGTGCTTAATGGGTAACTCATGTGTGACGCACTGAGGCTTAAAGATG 2775  
Db 2821 GGTTTTCTTGAACCTTGTGCTTAATGGGTAACTCATGTGTGACGCACTGAGGCTTAAAGATG 2880  
Qy 2776 AACCTGTGAAGACAAATGCGGTGAAGCTTGTGTGCGAGGTCAATGAGAAACATCTTGT 2835  
Db 2881 AACCTGTGAAGACAAATGCGGTGAAGCTTGTGTGCGAGGTCAATGAGAAACATCTTGT 2940  
Qy 2836 CATCTCAGAGACCTGAGTGTGTAAAGTGCATCTCACCGCGGCTGTGTGAGAGTCTG 2895  
Db 2941 CATCTCAGAGACCTGAGTGTGTAAAGTGCATCTCACCGCGGCTGTGTGAGAGTCTG 3000  
Qy 2896 TCTGTGTGATCTGAGAGAGACACCTGAAGAGCTGTGATCTCAAGCAATGCTCTG 2955

Db 3001 TCCTGTGTATCTCGAGAGACAGACACTGAAGAGCTGATCTCAACGACATGCCCCG 3060  
Qy 2956 GGTGACGGTGGGTTGCTGCGCTGTGCGGAGCTGAAGCAAAAGAAAGTGTCTGACG 3015  
Db 3061 GGTGACGGTGGGTTGCTGCGCTGTGCGGAGCTGAAGCAAAAGAAAGTGTCTGACG 3120  
Qy 3016 AGACTCGGGTTGAAGGACATGGAAGTCTGATCTGATGTTGTGAGGCACTCCCTGGCC 3075  
Db 3121 AGACTCGGGTTGAAGGACATGGAAGTCTGATCTGATGTTGTGAGGCACTCCCTGGCC 3180  
Qy 3076 CTTCCTGCAACCGGACATCTGAACAGTCTAAACCTGTGACAGAAATTAATCTTACGCCAA 3135  
Db 3181 CTTCCTGCAACCGGACATCTGAACAGTCTAAACCTGTGACAGAAATTAATCTTACGCCAA 3240  
Qy 3136 GGAATGATGAAGCTGTGTTGGCCCTTTGCTGCTGCTCCAGCTTAACCTTAAGATTAATGGG 3195  
Db 3241 GGAATGATGAAGCTGTGTTGGCCCTTTGCTGCTGCTCCAGCTTAACCTTAAGATTAATGGG 3300  
Qy 3196 CTGTGAAATGAGCAGTACCTGTCGAAATTAAGAAAGCTGTGAGGAGTGAAGTCACTC 3255  
Db 3301 CTGTGAAATGAGCAGTACCTGTCGAAATTAAGAAAGCTGTGAGGAGTGAAGTCACTC 3360  
Qy 3256 AAGCCCGAGTGTGATTAAGACGATGTTGGCATTCTTTTGATGAAGATGACCGAC 3312  
Db 3361 AAGCCCGAGTGTGATTAAGACGATGTTGGCATTCTTTTGATGAAGATGACCGGTAAC 3417

RESULT 7  
US-10-416-642-3  
; Sequence 3, Application US/10416642  
; Publication No. US20040043452A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: RAMKUMAR, Jayalaxmi  
; APPLICANT: ARVIZU, Chandra  
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0842 PCT  
; CURRENT APPLICATION NUMBER: US/10/416,642  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/249,407  
; PRIOR FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CB1  
US-10-416-642-3

Query Match 77.7%; Score 3133.6; DB 18; Length 3489;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

Qy 1 ATGAAGAGAGCAAAATCGCTCACTTTTCAGACTACGGGCTGCAATGCTGTCTATGAG 60  
Db 1 ATGAAGAGAGCAAAATCGCTCACTTTTCAGACTACGGGCTGCAATGCTGTCTATGAG 60  
Qy 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAATTAAGAAATTTCAAGAAATTC 120  
Db 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAATTAAGAAATTTCAAGAAATTC 120  
Qy 121 ACCAATGCTCTATTCACAGATTGAAATCGAAGTCCAACTGGAATGTCTGCACTC 180  
Db 121 ACCAATGCTCTATTCACAGATTGAAATCGAAGTCCAACTGGAATGTCTGCACTC 180  
Qy 181 CTCTTGATGATGATTAAGAGCATCGCTGGCTGAGCTTAAGCTTAAGCATTTTGA 240  
Db 181 CTCTTGATGATGATTAAGAGCATCGCTGGCTGAGCTTAAGCTTAAGCATTTTGA 240  
Qy 241 AAGATGAACCTGCAACCTCTCGAGAAAGGACAGGATGACATGAAAA----- 289

Db 241 AAGATGAACCTGCAACCTCTCGAGAAAGGACAGGATGACATGAAAAATTTACACAGA 300  
Qy 290 ----- 289  
Db 301 GATCTGAAGCAACGATGACTGACCAAGAACCAAGAAAGAAAGTCCAGAAAAATAA 360  
Qy 290 ----- 289  
Db 361 TATGGCATGACTAAGCTTATCTTGGGGTGTCTGACATCTGACTCGAATTAATTAAC 420  
Qy 290 ----- AAATTTCAAGCTATGAAACAAAGAT 318  
Db 421 AAGTATGTTGAATTCATTTCTTTTTCAGAAATTTTCAAGCTATGAAACAAAGAT 480  
Qy 319 GCCACAGCAGACAGACAGAAAGAAACAAAGAAATTTCAAGCTATGAAACAAAGATGTC 378  
Db 481 GCCACAGCAGACAGACAGAAAGAAACAAAGAAATTTCAAGCTATGAAACAAAGATGTC 540  
Qy 379 ACAGACAGACAGACAGAAAGAAACAAAGAAATTTCAAGCTATGAAACAAAGATGTC 438  
Db 541 ACAGACAGACAGACAGAAAGAAACAAAGAAATTTCAAGCTATGAAACAAAGATGTC 600  
Qy 439 GTGATGACCAATTCGCTGAGAGAGAGATGACGCTAGTTTGAAGAAACACTGCTCT 498  
Db 601 GTGATGACCAATTCGCTGAGAGAGAGATGACGCTAGTTTGAAGAAACACTGCTCT 660  
Qy 499 GACTGGCCGGAATGCAAAAGCTGTGCTGTGCTTTTGAATTCAGACCGTGGGCTTCCG 558  
Db 661 GACTGGCCGGAATGCAAAAGCTGTGCTGTGCTTTTGAATTCAGACCGTGGGCTTCCG 720  
Qy 559 CTGCGACAGGTTGTTCTGCAAGAAAGTCAAGAAATTTGGAAATTCGCTCTAGCCAGAG 618  
Db 721 CTGCGACAGGTTGTTCTGCAAGAAAGTCAAGAAATTTGGAAATTCGCTCTAGCCAGAG 780  
Qy 619 ATCGGCTGTGCTGGGCGCAAGTGGAGCTCTACAGGAAATGTTCTCTACGCTCTTTC 678  
Db 781 ATCGGCTGTGCTGGGCGCAAGTGGAGCTCTACAGGAAATGTTCTCTACGCTCTTTC 840  
Qy 679 CTCCCGCTTGAAGAGATGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738  
Db 841 CTCCCGCTTGAAGAGATGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Qy 739 GATGGCCAGACTCCAGGCTCCGCTGACGAGATCATGTCCGACAGAAAGCTGTTG 798  
Db 901 GATGGCCAGACTCCAGGCTCCGCTGACGAGATCATGTCCGACAGAAAGCTGTTG 960  
Qy 799 TTCAATCATTTGACGGTTGATGACCTGGGCTCTGCTCCCAATGACAAAGCTTGC 858  
Db 961 TTCAATCATTTGACGGTTGATGACCTGGGCTCTGCTCCCAATGACAAAGCTTGC 1020  
Qy 859 AAAGACTGGGCTGAGAAACAGCTCCGTTCAACCTCATACAGCTGCTGAGAGAGTTC 918  
Db 1021 AAAGACTGGGCTGAGAAACAGCTCCGTTCAACCTCATACAGCTGCTGAGAGAGTTC 1080  
Qy 919 CTGCTCCCTGAGTCTTCTGATGCTACCGCTCAGAGAGCTGGGACAGAGAGTCAAG 978  
Db 1081 CTGCTCCCTGAGTCTTCTGATGCTACCGCTCAGAGAGCTGGGACAGAGAGTCAAG 1140  
Qy 979 TCAGAGTGTGCTCCCGCTTACCTGTTAGTGAAGAAATCTCCGGGAAACAAAGATTC 1038  
Db 1141 TCAGAGTGTGCTCCCGCTTACCTGTTAGTGAAGAAATCTCCGGGAAACAAAGATTC 1200  
Qy 1039 CACTTGTCTTGTGACGCGGAGATTGTGAGCATGAGAGACAAAGGTTGCTGCTGATC 1098  
Db 1201 CACTTGTCTTGTGACGCGGAGATTGTGAGCATGAGAGACAAAGGTTGCTGCTGATC 1260  
Qy 1099 ATGAACAAACGCTGAGCTCTGACCAAGTCCAGGATGCTGCTGCTGCTGCTGCTGCT 1158  
Db 1261 ATGAACAAACGCTGAGCTCTGACCAAGTCCAGGATGCTGCTGCTGCTGCTGCTGCT 1320  
Qy 1159 GTGGCCCTGAGCTGAGAGAGTGTGGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218

Db 1321 GTGGCCCTGCAGCTGCAGAGCATGTGTGGGGAGAGCGTGCGCCCTTCAACCAACGCTC 1380  
Qy 1219 ACAGGCTGCAGCGCGCTTTGTGTGTTATGAGCTCAACCTCCAGAGGCTGTCCGGCC 1278  
Db 1381 ACAGGCTGCAGCGCGCTTTGTGTGTTATGAGCTCAACCTCCAGAGGCTGTCCGGCC 1440  
Qy 1279 TGTCTCAATCTGAGAGAAAGATTGTCTTGAAGCGCTTCTGCGATGTGTGTGAGGGA 1338  
Db 1441 TGTCTCAATCTGAGAGAAAGATTGTCTTGAAGCGCTTCTGCGATGTGTGTGAGGGA 1500  
Qy 1339 GTGTGAAATAGAAAGTCACTGTGTGTGATGATGCACTTCATGTGTCAAGATCTGGGAG 1398  
Db 1501 GTGTGAAATAGAAAGTCACTGTGTGTGATGATGCACTTCATGTGTCAAGATCTGGGAG 1560  
Qy 1399 TCTGAGCTCGGTCTGTCTTCAATGAAATTCCTTCCAGACAGCCACTGTGAGAG 1458  
Db 1561 TCTGAGCTCGGTCTGTCTTCAATGAAATTCCTTCCAGACAGCCACTGTGAGAG 1620  
Qy 1459 TACTACACCTTCTTCAACCTGATCTCCAGACTTGTGTGCGCGCTTGTACTACGTGTTA 1518  
Db 1621 TACTACACCTTCTTCAACCTGATCTCCAGACTTGTGTGCGCGCTTGTACTACGTGTTA 1680  
Qy 1519 GAGGCGCTGGAATAGAGCCAGCTCTGTCCCTCTGTACGTTGAGAGACAAAGAGTCC 1578  
Db 1681 GAGGCGCTGGAATAGAGCCAGCTCTGTCCCTCTGTACGTTGAGAGACAAAGAGTCC 1740  
Qy 1579 ATGAGCTTAAACAGGAGGCTTCCATATCACTCGCTTGGATGAAAGCGTTCTGTATT 1638  
Db 1741 ATGAGCTTAAACAGGAGGCTTCCATATCACTCGCTTGGATGAAAGCGTTCTGTATT 1800  
Qy 1639 GGCCTCTGAGCGAAGACGTAAAGAGGCACTGTAGAGTCTGTGCGGCTGTCCGTTCC 1698  
Db 1801 GGCCTCTGAGCGAAGACGTAAAGAGGCACTGTAGAGTCTGTGCGGCTGTCCGTTCC 1860  
Qy 1699 CTGGGGGTGAAGACGAACTTCTGCACTGGGTCTCTGTGTGGGTCAACACCTAATGCC 1758  
Db 1861 CTGGGGGTGAAGACGAACTTCTGCACTGGGTCTCTGTGTGGGTCAACACCTAATGCC 1920  
Qy 1759 ACCACCCAGAGACACCTGGAAGCGCTTCCACTGTCTTTTCAAGACTCAAGACAAAG 1818  
Db 1921 ACCACCCAGAGACACCTGGAAGCGCTTCCACTGTCTTTTCAAGACTCAAGACAAAG 1980  
Qy 1819 TTTGTTCCTTGCATTTAAAGCTTCCAAAGAGTGTGGCTTCCGATTTAAACAGAACTTG 1878  
Db 1981 TTTGTTCCTTGCATTTAAAGCTTCCAAAGAGTGTGGCTTCCGATTTAAACAGAACTTG 2040  
Qy 1879 GACTTGATAGCATCTTCTTCTGCTTCAAGCACTGTCCGATTTTGGGAAATTCGGGTG 1938  
Db 2041 GACTTGATAGCATCTTCTTCTGCTTCAAGCACTGTCCGATTTTGGGAAATTCGGGTG 2100  
Qy 1939 GATGTCAAAAGGATCTTCCCAAGAGATGATCCGCTGAGGACATGCTGTGCTCCCTCTA 1998  
Db 2101 GATGTCAAAAGGATCTTCCCAAGAGATGATCCGCTGAGGACATGCTGTGCTCCCTCTA 2160  
Qy 1999 TGTATGCGGATTAAGACCTCAATTGAGAGCATGTGGGAAATTTCTGCTCAATGCTTGC 2058  
Db 2161 TGTATGCGGATTAAGACCTCAATTGAGAGCATGTGGGAAATTTCTGCTCAATGCTTGC 2220  
Qy 2059 ACCCAACCAACCTGCGGAGCTGGAACCTGGGACAGACATCTTGAAGAGCGGCGCATG 2118  
Db 2221 ACCCAACCAACCTGCGGAGCTGGAACCTGGGACAGACATCTTGAAGAGCGGCGCATG 2280  
Qy 2119 AAGACCCGTGTGCCAAGCTGAGGACCTCACTGCAAGATTAAGAACCTCTATGTTTGA 2178  
Db 2281 AAGACCCGTGTGCCAAGCTGAGGACCTCACTGCAAGATTAAGAACCTCTATGTTTGA 2340  
Qy 2179 AATGACAGATTACCCCTGTGTGTGAGACCTCTGGAAGATGTGATGGCAACCGTAAC 2238  
Db 2341 AATGACAGATTACCCCTGTGTGTGAGACCTCTGGAAGATGTGATGGCAACCGTAAC 2400  
Qy 2239 CTAAAGATCCCTCAACTTGGAGAGCAACCACTGAAGAAAGATGTAAGATGGCGTGT 2298  
Db 2401 CTAAAGATCCCTCAACTTGGAGAGCAACCACTGAAGAAAGATGTAAGATGGCGTGT 2460

Qy 2299 GAACTTTAAAAACCCAAATGTTTGTGAGAGCTTTGAGGCTGATGTGCTGATTTG 2358  
Db 2461 GAACTTTAAAAACCCAAATGTTTGTGAGAGCTTTGAGGCTGATGTGCTGATTTG 2520  
Qy 2359 ACCCATGCTGTATCTGAAGATCTCCAAATCTTACGACCTTCCGAGCTGAAATCT 2418  
Db 2521 ACCCATGCTGTATCTGAAGATCTCCAAATCTTACGACCTTCCGAGCTGAAATCT 2580  
Qy 2419 CTGAGCTTGGCAGGAAACAAGGTGACAGACAGGAGATTAATGCTCTCAGTGAAGCTTG 2478  
Db 2581 CTGAGCTTGGCAGGAAACAAGGTGACAGACAGGAGATTAATGCTCTCAGTGAAGCTTG 2640  
Qy 2479 AGAGTCTCCAGTCCGCTGAGAGACTGATATCTGAGAGACTGTGACATCAACAGCAG 2538  
Db 2641 AGAGTCTCCAGTCCGCTGAGAGACTGATATCTGAGAGACTGTGACATCAACAGCAG 2700  
Qy 2539 GGTTCAGAGTCTGGCTTCAAGCTTGTGACGACCGGAGCTTGACACACTGTGCTTA 2598  
Db 2701 GGTTCAGAGTCTGGCTTCAAGCTTGTGACGACCGGAGCTTGACACACTGTGCTTA 2760  
Qy 2599 TCCAAACAAGCTTGGGAAAGAGGTGTAATCTATGTGTGATGCAATGAGGCTTCCC 2658  
Db 2761 TCCAAACAAGCTTGGGAAAGAGGTGTAATCTATGTGTGATGCAATGAGGCTTCCC 2820  
Qy 2659 CACTGTAGTCTGACAGAGCTGATGCTGAATCAAGTCCACCTGACACAGGCTGTGTGT 2718  
Db 2821 CACTGTAGTCTGACAGAGCTGATGCTGAATCAAGTCCACCTGACACAGGCTGTGTGT 2880  
Qy 2719 TTTCTTCACTTGGCTTAAAGGTAACTCAATGCTGACGACCTGAGCTTTAGATGAAAC 2778  
Db 2881 TTTCTTCACTTGGCTTAAAGGTAACTCAATGCTGACGACCTGAGCTTTAGATGAAAC 2940  
Qy 2779 CTGTGGAAGACATATGCGTGAAGCTTGTGCGAGGTCAATGAGAGACCATCTGTGCAT 2838  
Db 2941 CTGTGGAAGACATATGCGTGAAGCTTGTGCGAGGTCAATGAGAGACCATCTGTGCAT 3000  
Qy 2839 CTCAGAGCTTGGATGTAAAGTTCATCTCAACCCCGCTGTGTGAGAGTCTGTCC 2898  
Db 3001 CTCAGAGCTTGGATGTAAAGTTCATCTCAACCCCGCTGTGTGAGAGTCTGTCC 3060  
Qy 2899 TGTGTATCTTGAGAGACAGACCTGAAGAGCTGTGATCTCAAGCAATATGCCCTGGT 2958  
Db 3061 TGTGTATCTTGAGAGACAGACCTGAAGAGCTGTGATCTCAAGCAATATGCCCTGGT 3120  
Qy 2959 GACGCTGGGTGTGCTGGCTGTGCGAGGAGCTGAAGGAAAGAACATGTCTTGAAGAGA 3018  
Db 3121 GACGCTGGGTGTGCTGGCTGTGCGAGGAGCTGAAGGAAAGAACATGTCTTGAAGAGA 3180  
Qy 3019 CTGCGTGTGAAGGATGTGACTGATCTGATGCTGTGAGGACATCTCTTGGCCCTT 3078  
Db 3181 CTGCGTGTGAAGGATGTGACTGATCTGATGCTGTGAGGACATCTCTTGGCCCTT 3240  
Qy 3079 TCTTGAACCGGACATCTGACAGTCTTAAACCTGTGTGAAGATTAATCACTGCTCAAGAA 3138  
Db 3241 TCTTGAACCGGACATCTGACAGTCTTAAACCTGTGTGAAGATTAATCACTGCTCAAGAA 3300  
Qy 3139 ATGATGAAGCTGTGTGGCTTGTGCTGTCCACGCTTAACTTACAGATTAATGGGCTG 3198  
Db 3301 ATGATGAAGCTGTGTGGCTTGTGCTGTCCACGCTTAACTTACAGATTAATGGGCTG 3360  
Qy 3199 TGAAGATGGCAGTACCTGTGCAATTAAGAGAGCTGTGAGAGAGTGAAGTCACTCAAG 3258  
Db 3361 TGAAGATGGCAGTACCTGTGCAATTAAGAGAGCTGTGAGAGAGTGAAGTCACTCAAG 3420  
Qy 3359 CCCGAGTCTGTAATGAACGTATGTGCAATCTTTTGAATGAAGATGACGAC 3312  
Db 3421 CCCGAGTCTGTAATGAACGTATGTGCAATCTTTTGAATGAAGATGACGAC 3474

RESULT 8  
us-10-216-645-1  
; Sequence 1, Application US/10216645

Publication No. US20030125282A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
APPLICANT: LESSL, MONIKA  
APPLICANT: PETERS-KOTTIG, MICHAEL  
APPLICANT: BECKMANN, GEORG  
TITLE OF INVENTION: HUMAN MATER PROTEINS  
FILE REFERENCE: SCH-1910  
CURRENT APPLICATION NUMBER: US/10/216,645  
CURRENT FILING DATE: 2003-01-21  
PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 1  
LENGTH: 3926  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-216-645-1

Query Match 77.7%; Score 3133.6; DB 15; Length 3926;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

QY 1 ATGGAAGAGCAAAATCGCTACCTTTCCAGCTACGGGCTGCAATGGTCTCTATGAG 60  
DB 1 ATGGAAGAGCAAAATCGCTACCTTTCCAGCTACGGGCTGCAATGGTCTCTATGAG 60  
QY 61 CTAGACAAGAGAAATTTCAAGATTCAAGAAATTAAGAAATTTCTTCAAGATCG 120  
DB 61 CTAGACAAGAGAAATTTCAAGATTCAAGAAATTAAGAAATTTCTTCAAGATCG 120  
QY 121 ACCACATGCTCTATTCCACAGTTTGAATTCAGAAATGCCAAGTGTCTGCACTC 180  
DB 121 ACCACATGCTCTATTCCACAGTTTGAATTCAGAAATGCCAAGTGTCTGCACTC 180  
QY 181 CTTCTGATGATATTATGAGCATCGCTGCGCTGGGCTGCTCATTAAGATCTTTGAA 240  
DB 181 CTTCTGATGATATTATGAGCATCGCTGCGCTGGGCTGCTCATTAAGATCTTTGAA 240  
QY 241 AACATGAACCTGCGAACCTCTCGGAGAGGCAAGGATGACATGAAAA----- 289  
DB 241 AACATGAACCTGCGAACCTCTCGGAGAGGCAAGGATGACATGAAAA----- 289  
QY 290 ----- 289  
DB 290 ----- 289  
QY 301 GATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGAAAGATGCGAGAAAAATAA 360  
DB 301 GATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGAAAGATGCGAGAAAAATAA 360  
QY 290 ----- 289  
DB 290 ----- 289  
QY 361 TATGCGATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATATTAACAC 420  
DB 361 TATGCGATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATATTAACAC 420  
QY 290 ----- 289  
DB 290 ----- 289  
QY 421 AAGTATGTTGAATTCATTCTTTTTCAGAAATTTCAAGATGATGAAACAAAGAT 480  
DB 421 AAGTATGTTGAATTCATTCTTTTTCAGAAATTTCAAGATGATGAAACAAAGAT 480  
QY 319 GCCACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 378  
DB 319 GCCACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 378  
QY 481 GCCACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 540  
DB 481 GCCACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 540  
QY 379 ACAGACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 438  
DB 379 ACAGACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 438  
QY 541 ACAGACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 600  
DB 541 ACAGACAGACAGACAGAGAACAGAAATTTCAAGATGATGAAACAAAGATGATG 600  
QY 439 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 498  
DB 439 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 498  
QY 601 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 660  
DB 601 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 660  
QY 499 GACTGCGCGGAAATGCAAACTGCTGAGTCTTGTGATGATGATGATGATGATGATG 558  
DB 499 GACTGCGCGGAAATGCAAACTGCTGAGTCTTGTGATGATGATGATGATGATGATG 558  
QY 661 GACTGCGCGGAAATGCAAACTGCTGAGTCTTGTGATGATGATGATGATGATGATG 720  
DB 661 GACTGCGCGGAAATGCAAACTGCTGAGTCTTGTGATGATGATGATGATGATGATG 720  
QY 559 CCTGCAAGGTGTTCTGCAAGGAAATGCAAGAAATTTGGGAAATCGGCTCTAGCCAGAAAG 618  
DB 559 CCTGCAAGGTGTTCTGCAAGGAAATGCAAGAAATTTGGGAAATCGGCTCTAGCCAGAAAG 618

DB 721 CCTGCAAGGTGTTCTGCAAGGAAATGCAAGAAATTTGGGAAATCGGCTCTAGCCAGAAAG 780  
QY 619 ATGCTGCTGTGCTGGGCGCAAGGTGATCTTACAGAGAAATGTTCTCTAGCTCTTTC 678  
DB 781 ATGCTGCTGTGCTGGGCGCAAGGTGATCTTACAGAGAAATGTTCTCTAGCTCTTTC 840  
QY 679 CTCCCGGTTAGAGATGACAGCGGAAAGAGAGACATGTCACAGATTCATCTCCAG 738  
DB 841 CTCCCGGTTAGAGATGACAGCGGAAAGAGAGACATGTCACAGATTCATCTCCAG 900  
QY 739 GAGTGGCCAGACTCCCAAGCTCCGCTGACAGAGATCATGTCCTCCGACAGAAAGCTGTG 798  
DB 901 GAGTGGCCAGACTCCCAAGCTCCGCTGACAGAGATCATGTCCTCCGACAGAAAGCTGTG 960  
QY 799 TTCAATATTGACGTTTCTGATGACCTGGGCTCTGTCTCAACATGACAAAGCTCTGC 858  
DB 961 TTCAATATTGACGTTTCTGATGACCTGGGCTCTGTCTCAACATGACAAAGCTCTGC 1020  
QY 859 AAAAGCTGGGCTGAGAACGACCTCCGTTCAACCTCATAGCAAGTCTGAGAGAGTGC 918  
DB 1021 AAAAGCTGGGCTGAGAACGACCTCCGTTCAACCTCATAGCAAGTCTGAGAGAGTGC 1080  
QY 919 CTGCTCCCTGAGTCTCTGATGCTGACCGTCAAGAGAGTGGGCAAGAGAGCTCAAG 978  
DB 1081 CTGCTCCCTGAGTCTCTGATGCTGACCGTCAAGAGAGTGGGCAAGAGAGCTCAAG 1140  
QY 979 TCAAGGTGCTGTCTCCCGTTACCTGTTAGAGGAATCTCCGGGAAACAAAGATC 1038  
DB 1141 TCAAGGTGCTGTCTCCCGTTACCTGTTAGAGGAATCTCCGGGAAACAAAGATC 1200  
QY 1039 CACTTGTCTCTTGAAGCGGGGATTTGTGATGATCAAGAAACAAAGAGTTCCTGCTC 1098  
DB 1201 CACTTGTCTCTTGAAGCGGGGATTTGTGATGATCAAGAAACAAAGAGTTCCTGCTC 1260  
QY 1099 ATGAACCAACGAGTCTGCTGACAGAGTGGGAGGCGCGGCTGCTCTCATCTGC 1158  
DB 1261 ATGAACCAACGAGTCTGCTGACAGAGTGGGAGGCGCGGCTGCTCTCATCTGC 1320  
QY 1159 GTGGCCCTGCACTGCAAGAGTGTGGGAGAGAGCGTGCGCCCTTCAACCAACGCTC 1218  
DB 1321 GTGGCCCTGCACTGCAAGAGTGTGGGAGAGAGCGTGCGCCCTTCAACCAACGCTC 1380  
QY 1219 ACAAGCTCTGACAGCGCTTTTGTGTTCATGATGATCAACCTCTGAGAGGTGTCCGGG 1278  
DB 1381 ACAAGCTCTGACAGCGCTTTTGTGTTCATGATGATCAACCTCTGAGAGGTGTCCGGG 1440  
QY 1279 TGTCTCAATCTGAGAGAAAGATGCTCTGAAAGGCTTCTGCGATATGCTGTGAGAGGA 1338  
DB 1441 TGTCTCAATCTGAGAGAAAGATGCTCTGAAAGGCTTCTGCGATATGCTGTGAGAGGA 1500  
QY 1339 GTGTGAATTAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1398  
DB 1501 GTGTGAATTAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1560  
QY 1399 TCTGAGCTCGTCTCTGCTTTCATGATGATCAATCTTCTCCAGACAGCTGTGAGAG 1458  
DB 1561 TCTGAGCTCGTCTCTGCTTTCATGATGATCAATCTTCTCCAGACAGCTGTGAGAG 1620  
QY 1459 TACTACACCTTCTTCAACCTGATCTCCAGAGATCTTGTGCGGCTTGTACTAGTGTTA 1518  
DB 1621 TACTACACCTTCTTCAACCTGATCTCCAGAGATCTTGTGCGGCTTGTACTAGTGTTA 1680  
QY 1519 GAGGCGCTGAAATCGAGCCAGCTCTGCGCTCTGATGATGATGATGATGATGATG 1578  
DB 1681 GAGGCGCTGAAATCGAGCCAGCTCTGCGCTCTGATGATGATGATGATGATGATG 1740  
QY 1579 ATGAGCTTAAACAGGAGCTTCAATATCAATCTGCTTGTGATGATGATGATGATG 1638  
DB 1741 ATGAGCTTAAACAGGAGCTTCAATATCAATCTGCTTGTGATGATGATGATGATG 1800  
QY 1639 GGCCTGAGAGCAAGAGCTTAAGAGAGCACTGAGAGGCTGCTGAGGCTGCTCCGCTCC 1698  
DB 1800 GGCCTGAGAGCAAGAGCTTAAGAGAGCACTGAGAGGCTGCTGAGGCTGCTCCGCTCC 1698







APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: USSN 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: USSN 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: USSN 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: USSN 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: USSN 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: USSN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USSN 60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See file wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 347  
LENGTH: 3226  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (69)..(3168)  
US-10-092-900A-347

Query Match 71.9%; Score 2900.6; DB 18; Length 3226;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2503; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 406 CATGAGGTGACACATGGGACTACAGAGTCAGTGAATGACCAATTCGCTGAGGAGG 465  
DB 249 CATGAGGTGACACATGGGACTACAGAGTCAGTGAATGACCAATTCGCTGAGGAGG 308  
QY 466 GATGACGTCTGATGTTTGAAGAAACATGCTGCTGACCTGCGGAAATGCAACGTTGGCT 525  
DB 309 GATGACGTCTGATGTTTGAAGAAACATGCTGCTGACCTGCGGAAATGCAACGTTGGCT 368  
QY 526 GGTGCTTTTGAATTCAGACCCGCTGGGGCTTCCGGCTTCCGACCGGTGTTCTGACGGGAAG 585  
DB 369 GGTGCTTTTGAATTCAGACCCGCTGGGGCTTCCGGCTTCCGACCGGTGTTCTGACGGGAAG 428  
QY 586 TCAGGAATTTGGGAATTCGGCTCTAGCCAGAAAGATCGTGTGCTGAGGCGCAAGTGA 645  
DB 429 TCAGGAATTTGGGAATTCGGCTCTAGCCAGAAAGATCGTGTGCTGAGGCGCAAGTGA 488  
QY 646 CTCTACAGAGGAATGTTCTCTACGTTCTTCTTCTCCCGTTTGAAGATGACGGGAAG 705  
DB 489 CTCTACAGAGGAATGTTCTCTACGTTCTTCTTCTCCCGTTTGAAGATGACGGGAAG 548  
QY 706 AAGGAGACACGTGTCAACAGATTCATCTTCAGAGGATGCGGACATCCCAAGCTCCGGTG 765  
DB 549 AAGGAGACACGTGTCAACAGATTCATCTTCAGAGGATGCGGACATCCCAAGCTCCGGTG 608  
QY 766 ACGGATCATGTCCCGACACAGAAAGGCTGTGTTGTTTATCATTTGAGGTTGATGACCTG 825  
DB 609 ACGGATCATGTCCCGACACAGAAAGGCTGTGTTGTTTATCATTTGAGGTTGATGACCTG 668  
QY 826 GGTCTGTCTTCAACATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAAAGAGCTCTCCG 885  
DB 669 GGTCTGTCTTCAACATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAAAGAGCTCTCCG 728  
QY 886 TTCACCTCATACGAGATCTGCTGAGGAAGGTCTGCTCCCTGAGATCCTTCTGATCGTC 945

DB 729 TTCACCTCATACGAGATCTGCTGAGGAAGGTCTGCTCCCTGAGATCCTTCTGATCGTC 788  
QY 946 ACCGTCAAGACGTGGGCAACAGAAAGCTCAAGTCAAGAGGTGTCTTCCCGTTACTG 1005  
DB 789 ACCGTCAAGACGTGGGCAACAGAAAGCTCAAGTCAAGAGGTGTCTTCCCGTTACTG 848  
QY 1006 TTAGTTAAGGAATTCGGGGGAACAAGAAATTCACCTGCTCTTGAAGCGGGATTGGT 1065  
DB 849 TTAGTTAAGGAATTCGGGGGAACAAGAAATTCACCTGCTCTTGAAGCGGGATTGGT 908  
QY 1066 GAGCATCAGAAACACAAAGGTTGGTGCATCATGAAACAACCTGAGCTGCTCAGCAG 1125  
DB 909 GAGCATCAGAAACACAAAGGTTGGTGCATCATGAAACAACCTGAGCTGCTCAGCAG 968  
QY 1126 TCCAGGTCCCGCGGTGGGCTTCTCATCTGCGTGCCCTGACCTGACGACGTGGTG 1185  
DB 969 TCCAGGTCCCGCGGTGGGCTTCTCATCTGCGTGCCCTGACCTGACGACGTGGTG 1028  
QY 1186 GGGAGACGTCCTCCCTTCAACCAACGCTCAACAGGCTGACGCGGCTTTGTTGTT 1245  
DB 1029 GGGAGACGTCCTCCCTTCAACCAACGCTCAACAGGCTGACGCGGCTTTGTTGTT 1088  
QY 1246 CATGACCTCACCCCTCAGAGCGTGGTCCGGGCTGTCTCAATCTGAGAAAGATTGTC 1305  
DB 1089 CATGACCTCACCCCTCAGAGCGTGGTCCGGGCTGTCTCAATCTGAGAAAGATTGTC 1148  
QY 1306 CTGAAGCGCTTCTGCGGTATGAGTGTGAGGAGTGTGAATAGAACTCAAGTGTGAT 1365  
DB 1149 CTGAAGCGCTTCTGCGGTATGAGTGTGAGGAGTGTGAATAGAACTCAAGTGTGAT 1208  
QY 1366 GGTGACGACCTCATAGTTTCAAGAGACTGGGAGATCTGAGCTCCGCTCTGTTTCAATG 1425  
DB 1209 GGTGACGACCTCATAGTTTCAAGAGACTGGGAGATCTGAGCTCCGCTCTGTTTCAATG 1268  
QY 1426 AACATCTTCTCCAGACAGCCACTGTGAGGAGTACTACCTTCTTCACTCAGTCTC 1485  
DB 1269 AACATCTTCTCCAGACAGCCACTGTGAGGAGTACTACCTTCTTCACTCAGTCTC 1328  
QY 1486 CAGGACTTCTGTGCGGCTTGTACTACGTTTGAAGGGCTGGAATGACCGAGCTCTC 1545  
DB 1329 CAGGACTTCTGTGCGGCTTGTACTACGTTTGAAGGGCTGGAATGACCGAGCTCTC 1388  
QY 1546 TGCCTCTGTACGTTGAGAAACAAAGGTTCATGAGCTTTAAACAGGCTTCCAT 1605  
DB 1389 TGCCTCTGTACGTTGAGAAACAAAGGTTCATGAGCTTTAAACAGGCTTCCAT 1448  
QY 1606 ATTCACCTGCTTGGATGAAGGTTTCTGTTGGGCTCGTGAAGGAGATGTAAGAGG 1665  
DB 1449 ATTCACCTGCTTGGATGAAGGTTTCTGTTGGGCTCGTGAAGGAGATGTAAGAGG 1508  
QY 1666 CCACCTGAGGTCCTGCTGGGCTGTCCTTCCCTGCGGGTGAAGCAGAACTTCTGAC 1725  
DB 1509 CCACCTGAGGTCCTGCTGGGCTGTCCTTCCCTGCGGGTGAAGCAGAACTTCTGAC 1568  
QY 1726 TGGGCTCTGTGGGTCAAGAGCTTAATCCACACCCAGAGACACCTGAGCGCC 1785  
DB 1569 TGGGCTCTGTGGGTCAAGAGCTTAATCCACACCCAGAGACACCTGAGCGCC 1628  
QY 1786 TTTCACCTGCTTTCGAGACTCAAGACCAAGATTTGCTTGGCTTGAACCTTGA 1845  
DB 1629 TTTCACCTGCTTTCGAGACTCAAGACCAAGATTTGCTTGGCTTGAACCTTGA 1688  
QY 1846 CAGAAAGTGGGCTTCCGATTAAACAGAACTGACCTGATAGATCTTCTTGCCTC 1905  
DB 1689 CAGAAAGTGGGCTTCCGATTAAACAGAACTGACCTGATAGATCTTCTTGCCTC 1748  
QY 1906 CAGCACTGTCCGTAATTTGGGAAATTCGGGTGATGTCAAAAGGATTTTCCAAAGAT 1965  
DB 1749 CAGCACTGTCCGTAATTTGGGAAATTCGGGTGATGTCAAAAGGATTTTCCAAAGAT 1808  
QY 1966 GAGTCCGCTGAGGAGATGCTGCTGATGAGTCCCTTATGAGTGGGGAATAGACCTCATGAG 2025

Db 1809 GAGTCCGCTGAGGAGATGCTGTGGTCCCTCTATGATGCGGATGAGACCCCTCATTTGAG 1868  
Qy 2026 GAGCAGTGGGAAGATTTTGTGCTCAATGCTTGGACCCACCCACTGCGGAGCTGGAC 2085  
Db 1869 GAGCAGTGGGAAGATTTTGTGCTCAATGCTTGGACCCACCCACTGCGGAGCTGGAC 1928  
Qy 2086 CTGGGACAGCAGCATCTTGACAGAGCGGGCCATGAAAGACCTGTGTGCCAAGCTGAAGCAT 2145  
Db 1929 CTGGGACAGCAGCATCTTGACAGAGCGGGCCATGAAAGACCTGTGTGCCAAGCTGAAGCAT 1988  
Qy 2146 CCCACCTGCAAGATACAGACCCCTGATGTTTGAAGAAATGACAGATTACCCCTGTGTGAG 2205  
Db 1989 CCCACCTGCAAGATACAGACCCCTGATGTTTGAAGAAATGACAGATTACCCCTGTGTGAG 2048  
Qy 2206 CACCTGTGAGAAATGCTATGAGCCCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCC 2265  
Db 2049 CACCTGTGAGAAATGCTATGAGCCCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCC 2108  
Qy 2266 CACCTGAAAGAAAGAGATGTAAAGATGGCGGTGTAAAGCTTTAAACACCCAAATGTTTG 2335  
Db 2109 CACCTGAAAGAAAGAGATGTAAAGATGGCGGTGTAAAGCTTTAAACACCCAAATGTTTG 2168  
Qy 2326 TTGAGATCTTTGAGGCTGGATTTGCTGTGATTTGACCCATGCTCTTACTTGAAGATCTCC 2385  
Db 2169 TTGAGATCTTTGAGGCTGGATTTGCTGTGATTTGACCCATGCTCTTACTTGAAGATCTCC 2228  
Qy 2386 CAAATCTTACGACCTCCCGCAGCCTGAATCTGTAGACCTGGCAGAAACAGAGTGACA 2445  
Db 2229 CAAATCTTACGACCTCCCGCAGCCTGAATCTGTAGACCTGGCAGAAACAGAGTGACA 2288  
Qy 2446 GACCAAGGAGTAAATGCTCTCAGTGAATGCTTTGAGAGTCTCCAGTGGCCCTGAGAG 2505  
Db 2289 GACCAAGGAGTAAATGCTCTCAGTGAATGCTTTGAGAGTCTCCAGTGGCCCTGAGAG 2348  
Qy 2506 CTGATACCTGAGAGACTGTGGATACAGACCAAGGCTTGGCCAGAGTCTTGGCCCTC 2565  
Db 2349 CTGATACCTGAGAGACTGTGGATACAGACCAAGGCTTGGCCAGAGTCTTGGCCCTC 2408  
Qy 2566 GTGACCAACCGGAGCTTGACACCTGTGCTTATCCAAACAGGCTGGGAGAGAGAGT 2625  
Db 2409 GTGACCAACCGGAGCTTGACACCTGTGCTTATCCAAACAGGCTGGGAGAGAGAGT 2468  
Qy 2626 GTAAATCTACTGTGTGATGATGAGAGGCTTCCCACTGTAGTCTTGACAGAGCTGATGCTG 2685  
Db 2469 GTAAATCTACTGTGTGATGATGAGAGGCTTCCCACTGTAGTCTTGACAGAGCTGATGCTG 2528  
Qy 2686 AATCAGTGGCACTGAGACAGGCTGTGTGTGTTTCTTGACATTTGCCCTTATGGGTAC 2745  
Db 2529 AATCAGTGGCACTGAGACAGGCTGTGTGTGTTTCTTGACATTTGCCCTTATGGGTAC 2588  
Qy 2746 TCATGAGCTGACGCACTGAGCCTTATGACATGAAACCTGTGAAAGCAATGCGGTAAAGCTT 2805  
Db 2589 TCATGAGCTGACGCACTGAGCCTTATGAGCAATGAAACCTGTGAAAGCAATGCGGTAAAGCTT 2648  
Qy 2806 CTGTGCGAGGTCATGAGAGAACATCTTGTATCTCCAGACCTGAGTGTGTAAAGTGT 2865  
Db 2649 CTGTGCGAGGTCATGAGAGAACATCTTGTATCTCCAGACCTGAGTGTGTAAAGTGT 2708  
Qy 2866 CATCTACCCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTGAGAGACAGACACTG 2925  
Db 2709 CATCTACCCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTGAGAGACAGACACTG 2768  
Qy 2926 AAGAGCCTGATCTCAAGGACATGCTCTGGGTGACGGTGGGGTGTGCTGCGCTGTGACAG 2985  
Db 2769 AAGAGCCTGATCTCAAGGACATGCTCTGGGTGACGGTGGGGTGTGCTGCGCTGTGACAG 2828  
Qy 2986 GGAATGAGCAAAAGAACAGTGTCTGACAGACTCGGGTGTGAAGGCAATGTGTGACTGACT 3045  
Db 2829 GGAATGAGCAAAAGAACAGTGTCTGACAGACTCGGGTGTGAAGGCAATGTGTGACTGACT 2888  
Qy 3046 TCTGATGTGTGAGGCACTCTCTTGGCCCTTTCTTGCAACCGGACATCTGACAGTCTA 3105  
Db 2889 TCTGATGTGTGAGGCACTCTCTTGGCCCTTTCTTGCAACCGGACATCTGACAGTCTA 2948

Qy 3106 AACCTGTGAGATTAATCTTCACTGCCCAAGAGATGATGAGCTGTGTGGCCCTTTGGCC 3165  
Db 2949 AACCTGTGAGATTAATCTTCACTGCCCAAGAGATGATGAGCTGTGTGGCCCTTTGGCC 3008  
Qy 3166 TGTCCACGCTTAACTTACAGATTAATGGGCTGTGAAATGAGCACTACCTGTGCAATA 3225  
Db 3009 TGTCCACGCTTAACTTACAGATTAATGGGCTGTGAAATGAGCACTACCTGTGCAATA 3068  
Qy 3226 AGAAGCTGTGAGAGAGTGAAGCTTCAAGAGCCCGAGTGTATTTGACGGTATGG 3285  
Db 3069 AGAAGCTGTGAGAGAGTGAAGCTTCAAGAGCCCGAGTGTATTTGACGGTATGG 3128  
Qy 3286 CATCTTTTATGATGAGATGACCGACAC 3312  
Db 3129 CATCTTTTATGATGAGATGACCGGTAC 3155

RESULT 10  
US-10-399-443-5  
; Sequence 5, Application US/10399443  
; Publication No. US2004002869A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785  
; CURRENT APPLICATION NUMBER: US/10/399,443  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 3447  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-399-443-5

Query Match 31.7%; Score 1278; DB 17; Length 3447;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 908; Indels 18; Gaps 6;

Qy 498 TGACTGGCCGGAAATGCAAAAGTTGGCTGTGCTTTGATTGATTCAGACCGGTGGGCTTCGG 557  
Db 515 TGAAGCCCGCAAGATGAAATTAATGTCTGATGCTTTTAACATACAGAAAACCTTCCA 574  
Qy 558 GCCTCGACCGGTGTTCTGACCGGAAAGTCAAGAAATTTGGAAATTCGCTTAAGCAGAAG 617  
Db 575 GCTTCACACCAATTAATCTTACATGAGAACCGAGAGTTGGGAAGTCACTTTGGCAGAAG 634  
Qy 618 GATGCTGTGTGTGGGCGCAAGGTGACTTACCAAGGAATGTTCTCTACGCTCTT 677  
Db 635 TATGTTCTTGGCTGGGCAAGGTTAACTTCTCAAAAATG--TCTTTTGTCTATCTT 691  
Qy 678 CCTCCCGTTAGAGATGACGCGGAAGAGAGAGAGAGTGTCAAGAGTTTCACTCTCAG 737  
Db 692 CTTTCTGTTTGAAGAAATTAATGTTGACAGAGAGAGAGAGTTTGGACAGCTGATTTGCTTA 751  
Qy 738 GGAATGACAGACTCCAGGCTCCGTTGACGAGATCATGTCCGACAGAAAGCTGTT 797  
Db 752 GGAATGACAGACTCCGAGTCTTATGATCAAGATATCATGTCCCAACAGAAAGACTCTT 811  
Qy 798 GTTCATCATTTGACGTTTTCATGACCTGGGCTGTCTCTC--AAATGACAAAGCT 854  
Db 812 GTTGTATATGATGCTTGTGATATATGACTCTGTCTCTCAACATGATATATGACACT 871

QY	855	CTGC	AAAG	CTGG	GTG	AGG	AGC	AGC	CTC	CCG	TTCA	CCCT	CA	ACG	CA	AGC	CTG	CTA	GAA	914
Db	872	ATCC	GAGA	CTG	GA	AGAT	GA	GA	CA	GGCC	ATAT	AC	TCT	GAT	GTAC	AG	CTT	CA	GGA	931
QY	915	GATC	CTG	CTC	CTG	AGT	CC	TTT	CTG	AT	CGT	CA	CG	CT	CA	GAG	AC	G	AG	974
Db	932	GGCT	CTT	TA	ACT	CA	G	T	CTT	T	CT	AT	CA	T	T	CA	C	A	G	991
QY	975	CAAG	CTA	AGG	TC	GGT	CT	CCCC	GTTA	CT	GT	TAA	GGA	AT	CT	CC	GGG	AA	CA	103
Db	992	CAAG	CTA	AGG	TC	GGT	CT	CCCC	GTTA	CT	GT	TAA	GGA	AT	CT	CC	GGG	AA	CA	105
QY	1035	AAT	CA	CT	TG	CT	CC	CTT	CTG	AG	TC	GA	CT	CA	GA	GA	CA	CA	AG	109
Db	1052	ATCT	CAG	CTG	GT	CT	CT	CA	GA	AA	AT	CT	CA	A	T	GA	GT	CT	CA	111
QY	1095	GAT	CA	TGA	CA	AC	CGT	GA	CTG	CT	CA	CA	GT	CC	AG	GT	CC	CG	CT	115
Db	1112	TCTG	ATA	GA	AA	AT	CA	CA	AG	CT	GT	T	GA	CC	AA	TG	CA	AG	CC	117
QY	1155	CTG	CG	TGG	CC	CT	CA	G	CTG	CA	AG	TC	GA	AG	CG	T	GG	CC	CT	121
Db	1172	CTGT	BAG	GT	CT	T	CA	G	CTT	CA	BA	GA	AA	CT	GG	AA	B	GA	T	123
QY	1215	GCT	CA	CAG	CC	CT	CA	CG	CG	CT	T	T	GT	GT	T	AT	CA	AG	CT	127
Db	1232	TCT	CA	CC	GG	T	T	GT	AT	TG	CA	CG	TT	GT	T	CA	CA	AG	CT	129
QY	1275	GCG	CT	GT	CT	CA	T	CT	GA	GA	AA	AG	TT	GT	CT	GA	AG	CC	CT	133
Db	1292	GAG	CG	CT	CT	CA	G	GA	GA	AA	CA	GA	TTA	CT	T	AG	GG	TT	GT	135
QY	1335	GGA	GT	GA	GA	AT	GA	AG	AT	CA	GT	TT	GA	T	GT	GA	CA	CT	CA	139
Db	1352	AGA	GT	GA	GA	CA	GA	GT	CG	GT	TT	CT	A	T	A	T	GA	CT	GA	141
QY	1395	GGA	GT	GA	GA	CA	TC	TT	CT	CA	GA	CA	TC	TT	CT	CA	GA	CA	GC	145
Db	1412	GGA	GT	GA	GA	CA	TC	TT	CT	CA	GA	CA	TC	TT	CT	CA	GA	CA	GC	147
QY	1452	TGA	GA	GA	TA	CTA	CA	AC	CT	TT	CA	CA	CT	CA	GT	CA	GA	CA	CT	151
Db	1472	TGA	GA	GA	TA	CTA	CA	AC	CT	TT	CA	CA	CT	CA	GT	CA	GA	CA	CT	153
QY	1512	CGT	TTA	GAG	GG	CT	GA	AA	TC	GA	CA	CG	CT	CT	GA	CT	GA	GA	CA	157
Db	1532	TGT	TTA	GA	AG	GG	CT	GGA	--	GGA	T	GA	A	T	GA	CA	TT	TT	GG	158
QY	1572	GAG	GT	CA	T	GA	AG	CTT	AA	CA	GA	GG	CTT	CA	T	CA	CT	CG	CTT	163
Db	1589	GAG	CA	T	GA	AG	GG	CTT	AA	CA	GA	GG	CTT	CA	T	CA	CT	CG	CTT	164
QY	1632	CTT	GT	TT	GA	GG	CT	CG	TA	GA	GA	GA	CG	TA	GA	GG	CT	CG	TA	169
Db	1646	CTT	AT	TT	B	GG	CT	CA	T	GA	CA	GA	A	T	CTT	GA	AG	CT	GA	170
QY	1692	CGT	T	CC	CT	G	GG	GT	GA	AG	CA	GA	CTT	CT	GA	CT	G	GG	GT	175
Db	1706	CGT	AT	T	CC	CT	G	GG	GT	GA	AG	CA	CTT	CT	GA	CT	G			

OY	1932	TCGGGTGGAATCTCAAAGGAGATCTTCCAGAGAGATGAGTCGCGCTGAGGCAATGCTCGTGGT	1991
Db	1946	CCGGGTGGAATATCAAGAGCTCTCTCTCGGTAGATTAATCTTCGAGCTGTGCTCTGTGT	2005
OY	1992	CCCTCTAATGATGCGGGA---TAAGACCCCTATTGAGGAGCAGTGGGAGAAATTTCTGCTC	2048
Db	2006	TACTGTCCAGAGACACAAATGTAAAGCCCTCTCTCATGGAATGTGTGGGAAATCTTCTGCTC	2065
OY	2049	CATGCTTGGCAACCAACCACTGGCGAGCTGGACTGGGCGAGCAGCATCTGCACAGA	2108
Db	2066	TGTGCTTGGCAGCCTCCGGAATCTTGAAGAGAGCTGGACTTGGGGCGACAGCATCTGCAGTCA	2125
OY	2109	GCGGGCCATGAAAGACCTGTGTGTGCAAGCTGAGGCAATCCCACTGCAAGATATACAGACCT	2166
Db	2126	ACGGGCCATGAAAGATATCTGTGCTCGAGCTGGAAATCTCATGTCTGCACAATAATACGAAGCT	2185
OY	2169	GATGTTTAAATATCAACAGATTACCCCTGTGTGACAGCACTCTGGAGAAATGTCATGGC	2222
Db	2186	GACGTTTAAAGTGCAGAGTAAATGTGTCTGGCTGAACATCTCTGGAAGCTCTTTTAAAG	2244
OY	2229	CAACCGTAACCTTAAGATCTCCCTCAACTTGGAGGCAACCACTGAAGAAAGAGATGTAAG	2288
Db	2246	CAATCAAAACTTAAAGTACCTCAATCTAAGGAAACATCCCATGAAGGTATGATGACATGA	2305
OY	2289	GATGGCCGTGTAAACCTTTAAACACCCAAATATTTTGTGGAATCTTGAAGCTGTGATTG	2348
Db	2306	GTTAGCCCTGCACAAACGCTGAAACATCCAAATGTCTCGTGGAGACTGTGAAGTTGATTC	2365
OY	2349	CTGATGAATGACCAATGCTGTGTAACCTGAATCTCCCAATCTCTACGACTCCGCCAG	2408
Db	2366	CTGTGAATTAACCATCATTTGTTATGAGATATCTCCAGCTTTTATTTAACACACAG	2425
OY	2409	CCTGAATATCTTGAAGCTGTGCAGGAAACAAGGTGACAGACGAGAGATATGCTCTCAG	2468
Db	2426	GCTAAAGTGTCTGAGCCTGCGCACAATAATAGCTGGAGTAAAGATGATATCCTTGG	2485
OY	2469	TGATGCTTGAAGTCTTCCAGTGCCTGTGCAGAGCTGATATCTGAGAGCTGTGGCAT	2528
Db	2486	GAAATGCCCTTGAATGCTCAATATGTCTAATGCAAAAGTTGATATCTGCACACTGTGGCT	2545
OY	2529	CACAGCAACGGGTGGCAGAGTCTGGGCTGACGCCCTGTGACAAACGGAGACTTGACACA	2588
Db	2546	CACACTGGCAGCTGCACACTTCTGTGTCTCAGCCCTTTTACAGCAACAGAACTTGACACA	2605
OY	2589	CCTGTGCTTATCCAAACAACAGCTGTGGGAAACGAAGTGAATCTAATCTGTGTGATCCAT	2648
Db	2606	CTGTGCTGTCAACAACAACAGCTGTGGGAACTGAAGAGGTGCAACAGCTGTGTAGTTCTT	2665
OY	2649	GAGGCTTCCCACTGTAGTGTGCAAGAGCTGATGCTGAATCATGTGCCACTGACACGGC	2708
Db	2666	GAGGAATTCGAATATGTGCTCTCCAGCGGCTGATCTGAAATCATGTGCAACATTGTAGATGA	2725
OY	2709	TGCGTGTGTGTTTTTGTGCACTTGGGCTTATAGGGTAACTCATGTGCTGACGACTTGAGCTT	2766
Db	2726	TGCTTATAGGCTCTCTGGCAATGAGACTTGCACAAACAAGCTGACCCACTGAGCTT	2785
OY	2769	TAGCATGAACCTGTGGAAGACATATGGCGTGAACCTTGTGTCGAGGCTATGAGAAAC	2828
Db	2786	GACCATGAACCCGTAAGGAGTGTGTCAATGAACCTAATCTGTGTAAAGCTTTAAAGAAAC	2845
OY	2829	ATCTTGTATCTCCAGAGCTGTAGTTGTAAATGTCTATCTCAACGCGCGGTGCTGTGA	2888
Db	2846	TACTTGTATCTTCAAGAACCTGGAACCTAATGTGACTGTGCAACTGACACGAATCTGTGGA	2905
OY	2889	GAGTCTGTCTGTGTGATCTTGAGAGACACACTGAAGAGCTGTGATCTCACGACACA	2948
Db	2906	GGAATCTGACCTGTATGTATCAACACAAACAAAGCTTAAAGATTGATCTTGTGTAACA	2965
OY	2949	TGCGCTGGGTGAACGTGTGGGTGTGTGTGCGTGTGCGAGGAGCTGAAGCAAAAGACATGT	3008
Db	2966	CGCCCTGGGTGAACAAAGAGTCAATATCCCTGTGTGAAGGAGCTGAAGCAAAATGACAGCTC	3025
OY	3009	TCTGACGAGACTCGGGTTGAAGGCAATGACATGACTTCTGATTTGCTGTGAGGCACTCTC	3068

Db 3026 CCTGAGAGACTGTGGTTGGGGGCAATGTAAGTTGACTTCCAAATGCTGTGAGGCAATGTC 3085  
Qy 3089 CTGGCCCTTTCTCTGCAACCGGCACTGACAGTTAAACCTGTGCGAATTAATTGAG 3128  
Db 3086 ATGGCCATCTCTTGGCAACCCCTCACTTAACAGCTTAACCTGGGAAATGACTTGA 3145  
Qy 3129 TCCCAAGAAATGATGAAGCTGTGTGTCGGCCCTTGTGCTGTCCCAAGCTTAACATTA 3188  
Db 3146 TACATCGGGGATGTTGAAGCTGTGTCTGTGCTTCCAAATGCTGTCTTAACCTGGGAT 3205  
Qy 3189 AATTGGCTGTGGAATGCAAGTACCTGTGCAATTAAGAAAGCTGTGAGAAAGTGA 3248  
Db 3206 AATTGGCTGTGGAAGCAAGTACTATGCCGATGAAGAAGCAAGCTGAGAAAGTTGA 3265  
Qy 3249 GCTACTCAAGCCCGCAAGTCGTAATTGACGTAAGTTGGCAATCTTTGATGAAGATGACG 3308  
Db 3266 GTTGTCAAGCCCGCAAGTCGTAAGTGAATGTAATGCTAGTGAATGAAGATGACG 3325  
Qy 3309 ACAC 3312  
Db 3326 AAAC 3329

## RESULT 11

US-10-677-943-5

Sequence 5, Application US/10677943

Publication No. US20040072297A1

GENERAL INFORMATION:

APPLICANT: The Government of the United States of America as

APPLICANT: represented by the Secretary of the Department of Health and

APPLICANT: Human Services

APPLICANT: Nelson, Lawrence

APPLICANT: Tong, Zhi-Bin

TITLE OF INVENTION: Human Gene Critical to Fertility

FILE REFERENCE: 4239-64790

CURRENT APPLICATION NUMBER: US/10/677,943

CURRENT FILING DATE: 2003-10-01

PRIOR APPLICATION NUMBER: 60/241,510

PRIOR FILING DATE: 2000-10-18

PRIOR APPLICATION NUMBER: PCT/US02/09776

PRIOR FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: PCT/US01/10981

PRIOR FILING DATE: 2001-04-04

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn version 3.2

SEQ ID NO 5

LENGTH: 3447

TYPE: DNA

ORGANISM: Mus musculus

US-10-677-943-5

Query Match 31.7%; Score 1278; DB 18; Length 3447;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 908; Indels 18; Gaps 6;  
Qy 498 TGACGTGGCCGGAATGCAAAAGTGGCTGTGCTTTGATTCAGACGGTGGGGCTTCG 557  
Db 515 TGACGCCCCAGAGTAATTAATTTGTCTGATGCTTTTAAACCATACCAAAAACTTCA 574  
Qy 558 GCTTCGCAAGGTGTCTTTCGACGGAAGTCAAGATTGGAATGCGCTTACGCCAGAG 617  
Db 575 GCTTCGCAAGTATTCATCAATGAAAGCAAGATTGGAAGTCAAGCTTGGCCAGAG 634  
Qy 618 GATGCTGTGTGCTGGGCGCAAGTGAAGTCACTTACCAAGGAATGTTCTTACGTTCTT 677  
Db 635 TATGTTCTTGGCTGGGCAAGGTAACTCTTCCAAAAATG--TCCTTGTGATCTT 691  
Qy 678 CCTCCCGTTAGAGATGACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 737  
Db 692 CTCTCTTTAGAGAAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751  
Qy 738 GAAGTGGCCAGACTCCAGGCTCCGGTGAAGAGATCATGTCCGACCAAGAAAGCTGTT 797

Db 752 GAGAGTCCAGACTCTGGGATCTAGTGAAGAAATCATGTGCTCCCAACCAAGAAAGACTT 811  
Qy 798 GTTATCATTAAGCTTTTCCATGACCTGGGCTGTGTCTC---AACAATGACAAAAGCT 854  
Db 812 GTTGTCAATAGATGGCTTGGATGATGACTGTGTCTCCAAATATGATTAAGAACT 871  
Qy 855 CTGCAAGACTGGGCTGAGAAAGCAAGCTCCGTTCAACCTCAATAGCAAGTGTGAGGAA 914  
Db 872 ATCCAGAGACTGGAAGATGAACAGCCCATATATCTGATATGACAGCTCTGAGGAA 931  
Qy 915 GGTCTGTCTTGAAGTCTTCTGATGTGACCGTCAAGAGAGCTGGGCAAGAAAGCT 974  
Db 932 GGTCTCTTACCTCAGTCTCTTCTCATCATTAACCAAGCAAAACAGAGCTTAAGAAAAGT 991  
Qy 975 CAATGCAAGGTCTGTCTCCCGTTACCTTTAGTATGAGAGATCTCCGGGAAACAAG 1034  
Db 992 CAATGCAATGGTTGTGTCCTCTATATACTGGTTGAAGAGACTGTCTGCATCAAGAG 1051  
Qy 1035 AATCACTTGTCTCTTGAAGCGGGAATGGTGAAGCATCAGAAAGACAAAGGTTGCGTGC 1094  
Db 1052 ATCTCAGTGGTCTTCCAGAAACATCTCCATGAGTCTGATGAATACAAAGTCTCATTC 1111  
Qy 1095 GATCATGAACAACCGTGAAGCTGTGACCAAGTCCAGGTGCGCGGTGCTCTCAT 1154  
Db 1112 TCTGATGAATAATCACCAGCTGTTTGAACCAATGCGAGGCCCTGTGTGTGCTCCTGCT 1171  
Qy 1155 CTGCGTGGCCCTGAGAGTGAAGAGAGTGTGGGGAAGAGCGTCCGCCCTTCAACCAAG 1214  
Db 1172 CTGAGAGCTCTTAAGCTTACAGGAAGAACTGGGAAAGATGACCTTACCTTCCAGAC 1231  
Qy 1215 GCTCACAGGCTGTGACGCGCTTGTGATTCATCAGCTCAACCTCGAGGCGTGTGCTG 1274  
Db 1232 TCTACCGGTTTGTATGACCAAGTGTGTGTTTCAACCACTTCACTTGAAGAGCTTCCCA 1291  
Qy 1275 GCGCTGTCTCAATCTGGAAGAAAGATGTTCTTGAAGCGCTTCTGCGTATGCTGTGGA 1334  
Db 1292 GAGCGCTCTCAGTCAAGAAAGACAGATTAATCTAGTGGGTTGTGACATGAGAGCTGGA 1351  
Qy 1335 GGAAGTGTGAATAGAAAGTCAAGTGTGATGATGAGAGACTTCAATGTTCAAGAGACTCG 1394  
Db 1352 AGGAGTGTGACATGAGAGTGTGTTCTAAGATGAAGACCTGGAAGAACTATACCTTAAA 1411  
Qy 1395 GGAAGTGTGAGCTCGGTGCTGTGTTCAATGAATCATCTTCTCCAGACAGACC---ACTG 1451  
Db 1412 GAGTCTGAGATCTTGGCCCTCTTTCATGATGATCTTCTCAGAGTTGGCCACAAAG 1471  
Qy 1452 TGAGAGTACTACACTTCTTCCAGCTCAAGTCTCAGAGACTTGTGSCCGCTTGTACTA 1511  
Db 1472 TGACAGATGTTATGTTTCTCCACCTCAGGCTGAGAGATTTCTGTGCTTATATTA 1531  
Qy 1512 CGTGTGAAGGCGCTGGAATTCAGAGCAAGCTCTGCGCTCTGTACCTTGAAGAGACAA 1571  
Db 1532 TGTTTTGAAGAGGCTGGA---GGAATGAATACAGATTTTGTCTTCAAAAAACAAG 1588  
Qy 1572 GAGTTCATGAGAGCTTAAACAGGAGGCTTCCATATCCATCTGCTTGAAGAGCTTT 1631  
Db 1589 GAGCATATGAGAGGTGAAGAGAACTGACAGACTGCGCTC---CTCGAGATGAAGCGTTT 1645  
Qy 1632 CTGTGTGGCTGTGAGCGAAGAGTGAAGAGGCACTGAGAGGCTCTGTGGGCTGTCC 1691  
Db 1646 CTATTTTGGCTCATGAACAAGATATCTGAAAGACTTGAAGGTTCTGTTGAATATCC 1705  
Qy 1692 CGTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGTGGTCTCTGTGTGGGTCAAGAGCC 1751  
Db 1706 CGTATTTCAACTGTGTGAAGCAAGAGCTTCAACATCTGGGTCTCTGTGATGCTGACGAGT 1765  
Qy 1752 TAATGCAACCAAGCCAGAGAGCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1811  
Db 1766 CAATGCAAGAGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1825  
Qy 1812 CAAGAGTGTGTGCTTGGCATTTAAAGCTTCAAGAAAGTGTGCTTCGATTTAAACA 1871

1826 TGAAGATTGTTGGCGGGCTCTCAAGCGTTCAAGAGTGTGCTGATTACCA 1885  
1872 GAACCTGGAATTGATGATCTTCTTCCCTCCGACCTGCTCCGTAATTTGGGAAAT 1931  
1886 GAAAGATGGAATCTGAAAGCTCTCTTCTTCACTGTCTCAAGCACTGCAAACTTGAAGCAAT 1945  
1932 TGGGATGATGTCAAAGGATCTTCCCAAGATGATGATCGCTGAGGACATGCTCTGTGT 1991  
1946 CCGGGATGATTCAGAGACCTCTCTCGGTAGTAAATATCTCTGAGCTGTGCTGTGT 2005  
1992 CCGCTATGATGATGCGGGA--TAAACCTCATTTAGAGAGAGTGGAAATTTCTGTCT 2048  
2006 TACTGTCCAGAGACACAAATGTAAGCCCTCTCATAGAGTGTGGGAAACTTCTGTCTC 2065  
2049 CATGCTGGCAATCCCAACCACTGTGGGAGCTGTGACCTGTGGGAGAGATCTTGCACA 2108  
2066 TGTGCTGGCAGCTTCGGAATCTTGAAGAGCTGTGACTTGGGGAGAGATCTGTGATCA 2125  
2109 GCGGGCATGAAAGACCTGTGTGTCAGAGCTGAGGACCTCCACCTGCAAGATACAGACCT 2168  
2126 AGGGGCAATGAAGATATCTGTGCTGAGCTGCGGAATCTGCTGCAAGATACAGAGCT 2185  
2169 GATGTTAGAAATGCAAGATTAACCTGTGTGTGACACCTCTGAGAAATGCTCAATGCG 2228  
2186 GACGTTTAAAGTGAAGAGTGTGTCTGCGCTGAACAATCTCTGGAAGCTCTTTTGTAG 2245  
2229 CAACCTAACTTAAGATCCCTCACTTGGGAGGACCCACCTGAGAAAGAGATGTAG 2288  
2246 CAATCAAACTTAAGATCTCAATCTAGGAAACATCCCAAGAAAGATGATGATCAATGA 2305  
2289 GATGGGATGTAAGACCTTAACCAACCAAAATGTTGTTGATGATGAGGCTGATG 2348  
2306 GTTACGCTGGAAGGCTGAAACATCAAAATGCTCTGAGAGATCTTGAAGTTGATTC 2365  
2349 CTGTGATGTAAGCCATGCTCTTAACTCTGAAGATCTTCCAAATCTTACCACTCCCGAG 2408  
2366 CTGTGATTAACCATCATGTTTATGATGATCTCAGGCTTCTTATTTCAACCAAGCAG 2425  
2409 CTTGAATCTTGAAGCTGTGGCAGAAACAGATGACAGACAGGAGTAAATGCTCTCAG 2468  
2426 GCTAAAGTGTCTCAGCTGCGCAAAATGAGTGGAGTAAAGATATATCCCTTGTG 2485  
2469 TGAATGCTTGAAGTCTCCAGGCGCTGCAAGATGATATCTGAGAGCTGTGGCAT 2528  
2486 GATGCTTGAAGTCTCAATGCTGATCTGCAAAAGTGAATCTGGAACATCTGTGGCT 2545  
2529 CACAGCAAGGATGTCAGAGTCTGAGCTCAGGCTCTGACAGAACCGAGCTTGAACA 2588  
2546 CACACTGCGACGCTGCACTTCTGTGTCTCAGGCTCTTTCAGCAACAGAACTTGAACA 2605  
2589 CCTGTGCTATCCAAACAGAGCTGTGGGAAAGAGTGTAAATCTATGTGTGATCCAT 2648  
2606 CCTGTGCTGTCAAACAGAGCTGTGGGAACTGAAAGAGTGAACAGCTGTGCAATTCCT 2665  
2649 GAGGCTTCCCACTGATGTCAGAGGCTGATGCTGAATCACTGCGACCTGTGACAGCGC 2708  
2666 GAGGAATCAGAAATGCTCTCAGGCGCTGATATCTGAATCACTGAAATGTGTGATTA 2725  
2709 TGGGATGATGTTTCTTCACTTGGCGCTTATGAGTAACTGATGAGCTGAGCCTGAGCT 2768  
2726 TGTCTTATGCTCTCTGCGAATGAGCTTGAACAAACAGAAAGTGAACCTGAGCT 2785  
2769 TAGCATGAACCTGTGGAAGACATGAGCTTCTGTGTGAGAGTCAATGAGAAAC 2828  
2786 GACCAATGAACCCGTAGGGAGATGTCATGAAGCTATGTAAGCTTTTAAAGAAAC 2845  
2829 ATCTTGTCTATCTCCAGGACCTGAGATGTTGAAGTGTCTTCAACGCGCGGTGCTGTGA 2888  
2846 TACTGTGTTACTTTCAAGAACTGAAATGATGATGCTGCAACTCAACAGAACTGTGCGCA 2905  
2889 GAGTCTGTCTGTGATGATCTGAGAGAGACAGACCTGAAAGAGCTGTGATCTGACGGA 2948  
2906 GGAACCTGCGCTGATGATCAACAAACAGCACTTAAAGTTTGAATCTTGTGATCA 2965

2949 TGCCTGGGTGACGGTGGGTTGCTGCGCTGTGCGAGGGAATGAAGCAAAAGACATGT 3008  
2966 CGCCTTGGGTGACAAAGATGATTAACCTGTGTGAGGGAATGAAGCAAAAGTGAAGCTC 3025  
3009 TCTGACGAGACTCGGGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATG 3068  
3026 CTTGAGAGAGACTTGGGTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATG 3085  
3069 CTGGGCTCTTCTGCAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3128  
3086 ATTTGCAATCTTTCACACCTTCAACCTTCAACAGCTTAAACCTGTGAAGAAATGATTCAG 3145  
3129 TCCCAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3188  
3146 TACATGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3205  
3189 AATTGGGCTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3248  
3206 AATTGGCTGTGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3265  
3249 GCTACTCAAGCCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3308  
3266 GTTGTCAAGCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3325  
3309 ACAC 3312  
3326 AAC 3329

RESULT 12  
US-10-399-443-1  
; Sequence 1, Application US/10399443  
; Publication No. US20040028659A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785  
; CURRENT APPLICATION NUMBER: US/10/399,443  
; PRIOR FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-399-443-1

Query Match 28.6%; Score 1152.8; DB 17; Length 1157;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

754 CAGGCTCCGGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813  
1 CAGGCTCCGGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
814 TTGAT 873  
61 TTGAT 120  
874 AAGGAGCTCCGGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933  
121 AAGGAGCTCCGGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180



934 TTCTGATGTCACCGTCAGAGACGTGGGACACAGAAAGCTCAAGTCAGAGGTGCTCT 993  
Db TTCTGATGTCACCGTCAGAGACGTGGGACACAGAAAGCTCAAGTCAGAGGTGCTCT 240  
994 CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 1053  
Db CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 300  
994 CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 1053  
Db CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 300  
1054 CGCGGATTTGGTGGAGCATCAGAGACACAAAGGTTGGCGTGGATATGAACACCGTGGAG 1113  
Db CGCGGATTTGGTGGAGCATCAGAGACACAAAGGTTGGCGTGGATATGAACACCGTGGAG 360  
1114 CTGCTGACACAGTCGACAGTGGCCCGTGGGCTCTCATCTGCGGACCTGAGCTG 1173  
Db CTGCTGACACAGTCGACAGTGGCCCGTGGGCTCTCATCTGCGGACCTGAGCTG 420  
1174 CAGACGTGTGGGGAGAGCGTGGCCCTTCAACCAACGCTCAACAGGCTGACGCC 1233  
Db CAGACGTGTGGGGAGAGCGTGGCCCTTCAACCAACGCTCAACAGGCTGACGCC 480  
1234 GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGGTGGTCCGGCGTGTCTCATCTGGAG 1293  
Db GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGGTGGTCCGGCGTGTCTCATCTGGAG 540  
1294 GAAAGAGTTGTCTTGAAGCGCTTTCGCGGTATGGCTGTGAGAGGAGTGGATAGAG 1353  
Db GAAAGAGTTGTCTTGAAGCGCTTTCGCGGTATGGCTGTGAGAGGAGTGGATAGAG 600  
1354 TCAAGTTTGTATGTGACGACCTCATGTGTTCAAGAGACTCGGGAGTCTGAGCTCCGCT 1413  
Db TCAAGTTTGTATGTGACGACCTCATGTGTTCAAGAGACTCGGGAGTCTGAGCTCCGCT 660  
1414 CTGTTTCAATGAACATCTTCTCCGACACAGCCATGTGAGAGTACTAATCACTTCTTC 1473  
Db CTGTTTCAATGAACATCTTCTCCGACACAGCCATGTGAGAGTACTAATCACTTCTTC 720  
1474 CACCTCACTCTCAGAGACTTCTGTCGCCCTTGTACTACGTGTGAGAGGCTGGAAATC 1533  
Db CACCTCACTCTCAGAGACTTCTGTCGCCCTTGTACTACGTGTGAGAGGCTGGAAATC 780  
1534 GAGCCAGTCTCTGCTCTCTGTGACGTGAAGAAACAAGAGTCCATGAGCTTAAACAG 1593  
Db GAGCCAGTCTCTGCTCTCTGTGACGTGAAGAAACAAGAGTCCATGAGCTTAAACAG 840  
1594 GACGCTTCCATTCATCTGCTTGTGATGAAGGCTTCTGTTGGCTGCTGAGCGAA 1653  
Db GACGCTTCCATTCATCTGCTTGTGATGAAGGCTTCTGTTGGCTGCTGAGCGAA 900  
1654 GACGCTTCCATTCATCTGCTTGTGATGAAGGCTTCTGTTGGCTGCTGAGCGAA 1653  
Db GACGCTTCCATTCATCTGCTTGTGATGAAGGCTTCTGTTGGCTGCTGAGCGAA 900  
1714 AAGCTTGTGACATGAGGCTCTCTGTGTGGGTCAAGACCTTAATGCCACACCCAGAGAC 1773  
Db AAGCTTGTGACATGAGGCTCTCTGTGTGGGTCAAGACCTTAATGCCACACCCAGAGAC 1020  
1774 AACCCTGAGCGCTTCCATCTGCTTTCAGAGACTCAAGCAAAAGATTTGTTGCTGGCA 1833  
Db AACCCTGAGCGCTTCCATCTGCTTTCAGAGACTCAAGCAAAAGATTTGTTGCTGGCA 1080  
1834 TTAACACACTTCAAGAGAGTGTGCTTCCGATTAACAGAACTGGAATTGATGACT 1893  
Db TTAACACACTTCAAGAGAGTGTGCTTCCGATTAACAGAACTGGAATTGATGACT 1140  
1894 TCCTTCTGCTTCCAGC 1909  
Db TCCTTCTGCTTCCAGC 1156

RESULT 13  
US-10-677-943-1  
; Sequence 1, Application US/10677943

Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zhi-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT APPLICATION NUMBER: US/10/677,943  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US02/09776  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-677-943-1

Query Match 28.6%; Score 1152.8; DB 18; Length 1157;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

934 CAGGCTCCGGTGACGAGATCATGTCCGACACAGAAAGCTGTTGTTCAATCATTTAGCGGT 813  
Db CAAGCTCCGGTGACGAGATCATGTCCGACACAGAAAGCTGTTGTTCAATCATTTAGCGGT 60  
814 TTTCATGACCTGGGCTCTGTCTTCAACAAATGACCAAAAGCTCTGCAGAAAGCTGGGCTGAG 873  
Db TTTCATGACCTGGGCTCTGTCTTCAACAAATGACCAAAAGCTCTGCAGAAAGCTGGGCTGAG 120  
61 TTTCATGACCTGGGCTCTGTCTTCAACAAATGACCAAAAGCTCTGCAGAAAGCTGGGCTGAG 120  
874 AAGCAGCTCCGTTCAACCTCATACGAGTCTGTGAGAGGCTGCTCCCTGAGTCC 933  
Db AAGCAGCTCCGTTCAACCTCATACGAGTCTGTGAGAGGCTGCTCCCTGAGTCC 180  
121 AAGCAGCTCCGTTCAACCTCATACGAGTCTGTGAGAGGCTGCTCCCTGAGTCC 180  
934 TTCTGATGTCACCGTCAGAGACGTGGGACACAGAAAGCTCAAGTCAGAGGTGCTCT 993  
Db TTCTGATGTCACCGTCAGAGACGTGGGACACAGAAAGCTCAAGTCAGAGGTGCTCT 240  
181 TTCTGATGTCACCGTCAGAGACGTGGGACACAGAAAGCTCAAGTCAGAGGTGCTCT 240  
994 CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 1053  
Db CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 300  
241 CCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGATCACTGCTCTTGGAG 300  
1054 CGCGGATTTGGTGGAGCATCAGAGACACAAAGGTTGGCGTGGATATGAACACCGTGGAG 1113  
Db CGCGGATTTGGTGGAGCATCAGAGACACAAAGGTTGGCGTGGATATGAACACCGTGGAG 360  
301 CGCGGATTTGGTGGAGCATCAGAGACACAAAGGTTGGCGTGGATATGAACACCGTGGAG 360  
1114 CTGCTGACACAGTCGACAGTGGCCCGTGGGCTCTCATCTGCGGACCTGAGCTG 1173  
Db CTGCTGACACAGTCGACAGTGGCCCGTGGGCTCTCATCTGCGGACCTGAGCTG 420  
361 CTGCTGACACAGTCGACAGTGGCCCGTGGGCTCTCATCTGCGGACCTGAGCTG 420  
1174 CAGACGTGTGGGGAGAGGCTGCCCCCTTCAACCAACGCTCAACAGGCTGACGCC 1233  
Db CAGACGTGTGGGGAGAGGCTGCCCCCTTCAACCAACGCTCAACAGGCTGACGCC 480  
421 CAGACGTGTGGGGAGAGGCTGCCCCCTTCAACCAACGCTCAACAGGCTGACGCC 480  
1234 GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGGTGGTCCGGCGTGTCTCATCTGGAG 1293  
Db GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGGTGGTCCGGCGTGTCTCATCTGGAG 540  
481 GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGGTGGTCCGGCGTGTCTCATCTGGAG 540  
1294 GAAAGAGTTGTCTTGAAGCGCTTTCGCGGTATGGCTGTGAGAGGAGTGTGGAATAGAGAG 1353  
Db GAAAGAGTTGTCTTGAAGCGCTTTCGCGGTATGGCTGTGAGAGGAGTGTGGAATAGAGAG 600  
541 GAAAGAGTTGTCTTGAAGCGCTTTCGCGGTATGGCTGTGAGAGGAGTGTGGAATAGAGAG 600  
1354 TCAAGTTTGTATGTGACGACCTCATGTGTTCAAGAGACTCGGGAGTCTGAGCTCCGCT 1413  
Db TCAAGTTTGTATGTGACGACCTCATGTGTTCAAGAGACTCGGGAGTCTGAGCTCCGCT 660





TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-677-943-3

Query Match 18.8%; Score 756.6; DB 18; Length 1075;  
Best Local Similarity 99.5%; Pred. No. 66-218;  
Matches 759; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
OY 2550 TCTGGCCCTCAGCCCTCCGTCAGGAACCGGAGCTTGAACACCTGTGGCTATCCACAACAG 2609
Db 1 TCTGGCCCTCAGCCCTCCGTCAGGAACCGGAGCTTGAACACCTGTGGCTATCCACAACAG 60
OY 2610 CCTGGGGAACGAGGTAAATCTATCTGTGCATCCATGAGGCTTCCCACTGTAGTCT 2669
Db 61 CCTGGGGAACGAGGTAAATCTATCTGTGCATCCATGAGGCTTCCCACTGTAGTCT 120
OY 2670 GCAGAGCTGATGCTGAATCAAGTGCACCTGGAACGCGCTGGCTGTGATTTCTTGCACT 2729
Db 121 GCAGAGCTGATGCTGAATCAAGTGCACCTGGAACGCGCTGGCTGTGATTTCTTGCACT 180
OY 2730 TGGCCTTATGGTAACTCATGTGCTGACGCACTGAGCCTTAGCATGAACCTGTGGAAGA 2789
Db 181 TGGCCTTATGGTAACTCATGTGCTGACGCACTGAGCCTTAGCATGAACCTGTGGAAGA 240
OY 2790 CAATGGCGTGAAGCTTCTGTGCGAGGTATGAGAGAACCATCTTGTCTCAAGAACCT 2849
Db 241 CAATGGCGTGAAGCTTCTGTGCGAGGTATGAGAGAACCATCTTGTCTCAAGAACCT 300
OY 2850 GGAAGTTGATAAGTGTCACTTCACCGCCGCGTGTGAGAGTCTGCTGTGTATCTC 2909
Db 301 GGAAGTTGATAAGTGTCACTTCACCGCCGCGTGTGAGAGTCTGCTGTGTATCTC 360
OY 2910 GAGAGGCAACACCTGGAAGAGCTGTGATCTCACGCAATGCCCTGGGTGAAGTGGGT 2969
Db 361 GAGAGGCAACACCTGGAAGAGCTGTGATCTCACGCAATGCCCTGGGTGAAGTGGGT 420
OY 2970 TGCTGCGCTGTGCGAGGAGCTGAACCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAA 3029
Db 421 TGCTGCGCTGTGCGAGGAGCTGAACCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAA 480
OY 3030 GGCATGTGACTGACTTCTGATTTGTGAGGCACTCTTTGGCCCTTTCTCGCAACCG 3089
Db 481 GGCATGTGACTGACTTCTGATTTGTGAGGCACTCTTTGGCCCTTTCTCGCAACCG 540
OY 3090 GCATGTGACCAAGTCTAAACCTGTGCAAGATTAATTCACTGCCAAGGAATGAAAGCT 3149
Db 541 GCATGTGACCAAGTCTAAACCTGTGCAAGATTAATTCACTGCCAAGGAATGAAAGCT 600
OY 3150 GTGTTGCGGCTTTGCTGCCAGCTCTTAATTACAGATAATTGGGCTGTGAAATGGCA 3209
Db 601 GTGTTGCGGCTTTGCTGCCAGCTCTTAATTACAGATAATTGGGCTGTGAAATGGCA 660
OY 3210 GTACCTGTGCAATTAAGGAAGCTGTGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGT 3269
Db 661 GTACCTGTGCAATTAAGGAAGCTGTGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGT 720
OY 3270 AATTGACGTAAGTGGCATCTTTTGATGAAGATGACCGAAGC 3312
Db 721 AATTGACGTAAGTGGCATCTTTTGATGAAGATGACCGGTAAC 763
```

Search completed: July 19, 2005, 02:27:12  
Job time : 2360 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 16:09:10 ; Search time 12236 Seconds

(without alignments) 12552.250 Million cell updates/sec

Title: US-10-066-521-5

Sequence: 1 atggaagagagacaatcgct.....tcgaggcgctgctcctaa 4035

Scoring table: IDENTITY\_NUC

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1261	31.3	3475	3	AK087774 Mus muscu
2	723.8	11.9	791	7	CK000513 AGENCOURT
3	422.4	10.5	3359	3	BC021272 Homo sapi
4	422.4	10.5	3360	3	BC012789 Homo sapi
5	402.8	10.0	814	6	CB228956 AGENCOURT
6	381.4	9.5	776	7	CO814088 AGENCOURT
7	377.8	9.4	1913	3	AK016782 Mus muscu
8	374.4	9.3	808	7	CO804754 AGENCOURT
9	358.2	8.9	693	5	BU634350 AGENCOURT
10	357.2	8.9	684	5	BU630481 UI-H-FIL-
11	349.2	8.7	786	7	CO802335 AGENCOURT
12	336	8.3	792	7	CO82844 AGENCOURT
13	323.4	8.0	1119	4	BM454123 AGENCOURT
14	320	7.9	739	7	CO815646 AGENCOURT
15	313.6	7.8	642	5	BU618831 UI-H-FIL-
16	284.8	7.1	701	7	CO801395 AGENCOURT
17	282.4	7.0	817	7	CO797627 AGENCOURT
18	280.2	6.9	571	2	BB555431
19	276.6	6.9	3577	3	AK087843 Mus muscu
20	267.2	6.6	3218	3	AK054378 Mus muscu
21	256.2	6.3	666	4	AV367637 AV367637
22	247.4	6.1	579	1	BM228619 K0264H01-
23	242	6.0	565	4	BM226108 K0222C06-
24	238.2	5.9	589	6	CA559886 K0264H01-

25	232.6	5.8	552	7	CN724433	CN724433 E0879B11-
26	232.2	5.8	593	6	CA559979	CA559979 K0266B01-
27	231.2	5.7	593	6	CA558845	CA558845 K0250F09-
28	230.6	5.7	768	7	CO808001	CO808001 AGENCOURT
29	229.8	5.7	643	7	CF913803	CF913803 B0954D05-
30	224.8	5.6	636	7	CF914959	CF914959 B0971H02-
31	221.4	5.5	515	5	BM227562	BM227562 K0245E10-
32	221.4	5.5	4123	3	HSMB00983	AL117470 Homo sapi
33	221	5.5	511	4	BM230035	BM230035 K0288H05-
34	221	5.5	512	4	BM226240	BM226240 K0258A02-
35	219.8	5.4	513	4	BM227869	BM227869 K0250F09-
36	218.6	5.4	519	4	BG071729	BG071729 H3102E07-
37	218.6	5.4	519	4	BG084574	BG084574 H3102E07-
38	217.8	5.4	559	6	CA561556	CA561556 K0288H05-
39	216.2	5.4	743	7	CK35847	CK35847 UI-M-HNO-
40	213.8	5.3	3275	3	AK054426	AK054426 Mus muscu
41	209.8	5.2	599	7	CF915731	CF915731 B0984E03-
42	203.8	5.1	473	4	BM227899	BM227899 K0251B09-
43	203.4	5.0	1434	3	AK014932	AK014932 Mus muscu
44	201.2	5.0	1417	3	CF594199	CF594199 full-leng
45	201.2	5.0	1586	3	CR607466	CR607466 full-leng

## ALIGNMENTS

RESULT 1	AK087774	LOCUS	AK087774	3475 bp	mRNA	linear	HTC 03-APR-2004
DEFINITION	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E830019P16 product:MUSE MATER protein (maternal-antigen-chat-embryos-require) protein, full insert sequence.						
ACCESSION	AK087774	GI:26104500					
VERSION	AK087774.1						
KEYWORDS	HTC; CAP trapper.						
SOURCE	Mus musculus (house mouse)						
ORGANISM	Mus musculus						
REFERENCE	1						
AUTHORS	Carninci, P. and Hayashizaki, Y.						
TITLE	High-efficiency full-length cDNA cloning						
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)						
MEDLINE	99279253						
PUBMED	10349636						
REFERENCE	2						
AUTHORS	Garnici, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.						
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes						
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)						
MEDLINE	20499374						
PUBMED	11042159						
REFERENCE	3						
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.						
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer						
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)						
MEDLINE	20530913						
PUBMED	11076861						
REFERENCE	4						
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium						
TITLE	Functional annotation of a full-length mouse cDNA collection						
JOURNAL	Nature 409, 685-690 (2001)						
REFERENCE	5						

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3475)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanezaki,T., Haru,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Horii,F., Imotani,K., Ishii,Y., Itchi,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazune,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sugabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-Apr-2002) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), Riken Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gscc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. <a href="http://genome.gsc.riken.jp/">URL:http://genome.gsc.riken.jp/</a> . <a href="http://fantom.genic.riken.jp/">URL:http://fantom.genic.riken.jp/</a> .
FEATURES	location/Qualifiers source 1..3475 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:E330019P16" /db_xref="taxon:10090" /clone="E3j0019P16" /sex="female" /tissue_type="ovary" /clone_lib="RIKEN full-length enriched mouse cdna library" /dev_stage="2 days pregnant adult" 77.._3356 /note="mouse MATER protein (maternal-antigen-that-embryos-require) protein putative" 3452..3457 /note="putative" 3475 /note="putative"
ORIGIN	
Query Match	31.3%; Score 1261; DB 3; Length 3475;
Best Local Similarity	65.3%; Pred. No. 0;
Matches 1982; Conservative	0; Mismatches 1035; Indels 20; Gaps 8
OY	287 AAAAAATTTCACAAAGCTATGGAAACAAGAAGTGCCACACGACAGACAGAACAGAAACAAG 346 
Dd	324 AAAATGTCACTAGAGCAATCTCGAAGAAGACAGTGAATCAGAGAAAGTGAACAGGGGTGAG 383 
OY	347 AAATTTCACAAAGCTATGGAAACAAGAAGTGCCACACGACAGACAGAACAGAAACAAGAC 406 
Dd	384 AAAGAAAATGACTTTCTCCAGAAAACGACAGTAATCAATCCAGAAAAGCCAAGACGAG 443 
OY	407 ATGAGAGTGACCATGAGGACTCTCAAGAGTCACTGATGACCAAATTCGTGAGAGGAGAG 466 
Dd	444 AGCAGAGCACACGATCAAGTTAATGGAGGTGACTTCAACAGACTCAACAGGCCCATGTGATTG 503 
OY	467 ATGTAAGTCGTAGTTTTGAAAAACATGCTGCTGCAGCTGGCCGAGAAATGCAAGCTTGCTG 526

Db	504	CTAAGTTCGACACAAgTGTGATCTACACTATGACAGGCCAGAGATGMAATTAATGTCG	563
Oy	527	GTGCTTTTGAATTCAGACCCGTGGGGCTTCCGGCCCTGCACGGTGGTTCTGCACGGAAAGT	586
Db	554	ATGCTTTTAAACATACCAAGAAAACCTTCCAGCTCCACACATATCTCAATGGAAAGC	623
Oy	587	CAGGAATTTGGAA-ATCGGCTCTAGCCAGANAGATCGTGTGTGCTGGGCGCAAGTGA	645
Db	624	CAGGAGTTGGGAAGCTCAGCTTTGGCCAGAAAGTATTGTTCTTGCTGGGCGCACAGGCTAA	683
Oy	646	CTTATCCAGGGAAATTTCTCCTAAGCTTTCTTCTCCCGCTTAAGAGATGCACGGAG	705
Db	684	CTCTTCCAAAAAATG--TCTTTTGTCATCTTCTTCTGTGAAGAAATTAAGTGA	740
Oy	706	AAGGAGACAGTGTTCACAGAG-TTCATCTCAGAGAGTGGCACAATCCAGGCTCCGGT	764
Db	741	GAGAAAGCAATTTGGCACAGCTGATTTGCTAAGAGTGTCCAACTCTGGAGCTTAGT	800
Oy	765	GACGAGATCATGTCTCCGACCAAGAGGCTGTGTTCATCATTTAGCGTTTCATGACCT	824
Db	801	GACAAAGATCATGTCTCCACAGAAAGCTCTGGTTGTCTATGATGGCTTGATGATAT	860
Oy	825	GGGCTTGTCTCTC--AACATGACCAAAGCTCTGCAAAACATGGGCTGAGAAAGCAGCC	881
Db	861	GGACTGTCTCTCCAACTGATGATATGACACTATCCAGAACCTGGAAGATGAACAGCC	920
Oy	882	TCCGTTACCCCTCATACGCACTGTGCTGAGAGAGTCTGTCTCCCTGAGTCTCTTCGAT	941
Db	921	CATATACATCTGTATGTACAGCTCTCTGAGAAAGGCTCTCTACTCTAGTCTTTCTCAT	980
Oy	942	CGTCAACGTCAGAGACGTGGGCACAGAGAGCTCAAGTCAGAGGTCTGTCTCCCGTTA	1001
Db	981	CATTACCAACAGAAACAGAGCTTAGAAAACTCAAGTCATGGTGTGTCTCCCTCTA	1048
Oy	1002	CCTGTATGTAGAGAAATCTCCGGGGAACAAAGATTCATCTTGTCTCTTGAAGCGGGAT	1061
Db	1041	TATATCTGTTTAAGAGACTGTCTGATCAAGAGATCTCAGCTGTCTCCGAAATCTC	1100
Oy	1062	TGGTAGATCAGAGAACACAAAGGTTGCGTGCATGATGAACAAAGCTGAGCTGCTGA	1121
Db	1101	CAATGATCTATATGAAATACAAATGCTTTCATTTCTGTATGAATAATCACAGCTGTTGA	1160
Oy	1122	CCAGTGCAGAGTGCCTGCGGCTGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTCAGAGC	1181
Db	1161	CCAAATGCCAGGCCCTCTGTGTGTCTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1220
Oy	1182	GGTGGGGAGAGCGCTGCGCCCTTCACAAACGCTCACAGGCTGCAAGCGCTTTGT	1241
Db	1221	ACTGGAAAGAGATGCACCTTACCTGCGACACTCTCACCGTGTGTATGCTCAAGTTGGT	1280
Oy	1242	GTTTCATGAGTCACCCCTGAGGGCGTGTGCGGCGCTCTCATATCGAGGAAAGAT	1301
Db	1281	GTTCACCAAGTCACTTTGAAAGAGCCTTCCCAAGGCGCTTCAGTCAGAGAAACAGAT	1340
Oy	1302	TGTCCTGAAGCGCTTGTGCGCTATGAGCTGTGAGAGGAGTGTGAAATAGAAATCAGTGT	1361
Db	1341	TACTCTAGTGGGTTTGTGTGATGATGAGCAGCTGAAAGAGTGTGACATGAGGTGGTGT	1400
Oy	1362	TGATGTGACGACCTCATGTTCAAGGACTCGGGAAGTCTGAGCTCCGTCTGTGTCA	1421
Db	1401	CTATGATGATGACCTGAAAGAACTATAGGCTTAAGAGAGTCTGAGATCTTGCGCTCTTCA	1460
Oy	1422	CATGAACATCTTTTCTCCAGCAACCC--ACTGTGAAGAGTACATACATCTTTCCACT	1478
Db	1461	CATGAACATCTTTTCTCCAGGTTGGCCACAAAGTAGAGAGTATATGTTTCTCCACT	1520
Oy	1479	CAGTCTCAGAGACTTGTGTGCGGCTTGATCTAGTGTAGAGGCTCGAAATTCGAGCC	1538
Db	1521	CAGCTCGAGAGATTTCTTGTCTGCTTATATTAATGTTTGAAGAGGCTTGA--GGAATG	1577
Oy	1539	AGCTCTTGCCTCTGTATCGTTGAAGACAAAGAGTCCATGAGCTTTAAACAGCAGG	1598

Db 1578 GAATCAGCATTTTGGCTTCATTGAAAAACCAAGAGCATCATGAGGTGAAGAGAACTGA 1637  
 Qy 1599 CTTCCATATCCACTGGCTTTGGAAGAAGGTTTCTGTTTGGCTTCGAGGAGAGACT 1658  
 Db 1638 CGACACTGGCTTC---CTCGGAGTGAAGGTTTCTTATTTGGCTTCATGAACAGAAATAT 1694  
 Qy 1659 AAGAGGCACTGAGAGTCTGCTGGGCTGTCCCGTTCCCTGGGGGGTGAACAAAGT 1718  
 Db 1655 CTGGAAGACTGGAAGTCTCTTTGAAATATCCCGTATTCACACTGTTGAGCAAGAGT 1754  
 Qy 1719 TCTGCACTGGGTCTCTCTGTTGGGTCAAGAGCTTAATGCCACACCCCAAGAGACCT 1778  
 Db 1755 CCAACACTGGGTCTCTGATAGCTCAGAGGTCAATGGCAACAGCCCAATGACACCT 1814  
 Qy 1779 GGAACCTTCCACTGTCTTTTGAAGCTCAAGACAAAGATTGTTCCGCTTGGCAATTAA 1838  
 Db 1815 GGATTCCTTCTATTTCTATTGAGTCTCAGATGAAGATTGTTGGGGGGCTCTCA 1874  
 Qy 1839 CAGCTTCAAGAGTGGCTTCGATTAACAGAACTGACCTTGATAGACTTTCCCT 1898  
 Db 1875 ACGCTTCAAGAGTGGCTTCGATTAACAGAACTGAGCTTGAGAGTCTTCTCA 1934  
 Qy 1899 CTGCTTCAAGACTGTCCGATTTTGGGAAATTCGGGTGATGTCAAGAGGATCTTCC 1958  
 Db 1935 CTGCTTCAAGACTGTCAAGAACTTAAGGCAATCCGGGTGATTAACAGACTCTCTC 1994  
 Qy 1959 AAGAGTATGTCCTGAGGCAATGCTGTGCTCTCTGATG--GATCGGGAATAGAC 2015  
 Db 1995 GGTATATATATCTCTGAGCTGTGCTGTGTTACTGTCCGGAGACCAATGTAAAGC 2054  
 Qy 2016 CCTCATGAGAGAGTGGAGATTTTCTGCTCCAGCTTGGCAACCCCAACCTGCG 2075  
 Db 2055 CTTCTCATGAGTGGTGGGAACTCTGCTGTGCTGCTGGCACTTCGGAACTTGA 2114  
 Qy 2076 GCAAGTGAAGCTGGGAGAGCATCTCTGACAGAGCGGCAATGAAGACCTGTGCGCA 2135  
 Db 2115 GGAAGTGAAGCTGGGAGAGCATCTCTGACAGAGCGGCAATGAAGATCTGTGCTGA 2174  
 Qy 2136 GCTGAGGCACTCCCTGCAAGATACAGACCTGATGTTTGAAGATGCAAGATTAACCC 2195  
 Db 2175 GCTGAGGATAGTCTGCAAGATACAGAGCTGACGTTAAGATGCAAGAGTGTGTC 2234  
 Qy 2196 TGTGTCGAGCACTCTGAGAAATGCTCATGCGCAACCTTACCTAAGATCCCTAACT 2255  
 Db 2235 TGGCTTGAAGCACTCTGAGAGCTCTTTTGAATCAAACTTAAGTACTTCAACT 2294  
 Qy 2256 GGGAGGCAACCCACTGAGAGAAAGATGATGAGTGGGTGAGAGCTTAAACACC 2315  
 Db 2295 AGGGAACACTCCCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 2354  
 Qy 2316 AAAATGTTGTTGAGTCTTTGAGGCTGATTTGCTGTGATTTGACCCATGCTTACT 2375  
 Db 2355 AAAAGTCTCCGTTGAGAGCTCTGAGGTGGAATTCGTGTGATTAACATCATTTGTTA 2414  
 Qy 2376 GAAATATCTCCAAATCTTACGACTCCCGCAAGCTGAAATCTCTGAGCTTGGCAAGAA 2435  
 Db 2415 GATGATCTTCCAGCTTCTTATTTCAACCAACGAGTAAAGTCTTCAAGCTTGGCAAAA 2474  
 Qy 2436 CAAGGTGACAGACGAGGAGTATGCTCTGATGATGCTTGAAGTCTCCAGTGGC 2495  
 Db 2475 TGAAGTGGAGTAAAGATATATCTCTTGGAAATGCTTGAAGTCTCAATGTCT 2534  
 Qy 2496 CTTGAGAAAGTATGATGAGAGTGTGAGCTCAAGCAGCGGTGTCAGAGTCTGGC 2555  
 Db 2535 ACTGCAAAAGTTGATGATGAGCACTGTGCTCAACCTGCACTGCACTTCTGAT 2594  
 Qy 2556 CTCACCCCTCTGACCAACCGGAGCTTGAACACTGTGCTTATCCAAACAGCTGGG 2615  
 Db 2595 CTCACCCCTCTTACCAACCAACCTTGAACACTGTGCTTATCCAAACAGCTGGG 2654  
 Qy 2616 GAAACAAGGTATATCTACTGTGTGATTCATGAGGCTTCCCACTGATGCTGAGAG 2675  
 Db 2655 GACTGAAGAGTGAACAGCTGTGTGATGCTTGAAGAAATCAAGATGTGCTCTCAAGC 2714

Qy 2676 GGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2735  
 Db 2715 GGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2774  
 Qy 2736 TATGAGTAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2795  
 Db 2775 TGCAAAAGCAACAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2834  
 Qy 2796 CGTGAAGCTTCTGTCGAGTCAATGAGAAACCATCTTGTATCTTCAAGACTGAGAGT 2855  
 Db 2835 AATGAAGCTTACTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2894  
 Qy 2856 GGTAAAGTATCTCAACCGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2915  
 Db 2895 AGTGAAGCTCAACCTCAACAGAACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2954  
 Qy 2916 CAGACACCTGAAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2975  
 Db 2955 CAGACCTTAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3014  
 Qy 2976 GCTGTGAGGAGTGAAGCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3035  
 Db 3015 CTTGTGAGGAGTGAAGCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3074  
 Qy 3036 TGAAGTATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3095  
 Db 3075 TAAAGTATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3134  
 Qy 3096 GACCACTTAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3155  
 Db 3135 GACCACTTAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3194  
 Qy 3156 GACCTTTCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3215  
 Db 3195 TGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3254  
 Qy 3216 TGTGAAATTAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3275  
 Db 3255 TGGCAGTGTGAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3314  
 Qy 3276 CGGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3312  
 Db 3315 TGTGATG 3351

RESULT 2  
 CR000513  
 LOCUS 791 bp mRNA linear EST 26-NOV-2003  
 DEFINITION AGENCOURT\_16368905 NIH\_MGC\_221 Homo sapiens cDNA clone  
 IMAGE:30708637 5', mRNA sequence.  
 ACCESSION CR000513  
 VERSION CR000513.1 GI:38526547  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 791)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 AUTHORS Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NIH  
 National Cancer Institute  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgs@bbs-remail.nih.gov  
 Tissue Procurement: James Martin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:

http://image.llnl.gov  
 Plate: NDAMI075 row: c column: 14  
 High quality sequence stop: 688.  
 Location/Qualifiers

# FEATURES

source

```
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708637"
/lab_host="DH10B Tona"
/clone_id="NTH_MGC_221"
/note="Organ: mixed; Vector: pYX-Ase; Site_1: EcoRI;
Site_2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Ase vector. Average insert size
4-5kb. Adaptors 5' (AATTCGACGAGG) 3' and 5' d
(CCTGTCGCG) 3'. 3' linker sequence - GCGGCGGTGAGAGCC T18.
Sequencing primers 3' end: T3 promoter primer 5'd
(AATACCTCTACTTAAGGA) 3'. 5' End: T7 promoter primer 5'd
(TAATGACCTACTTAAGG) 3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NTH_MGC
Library"
```

## ORIGIN

```
Query Match      17.9%; Score 723.8; DB 7; Length 791;
Best Local Similarity 99.1%; Pred. No. 1.6e-188;
Matches 780; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Oy 565 ACGGTTCTTGACGAGAAAGTCAGAGAAATGGGCTCTAGCCAGAGATCGTG 624
Db 5 AGGGTCTTTCGACGAGAAATGGGAAATGGCTTACCGAAGATCTGTG 64
Oy 625 CTGTGCTGGGCGCAAGTGAAGTCTTACCGAGGAATGTTCTCTTCTTCTCC 684
Db 65 CTGTGCTGGGCGCAAGTGAAGTCTTACCGAGGAATGTTCTCTTCTTCTCC 124
Oy 685 GTTAGAGATGACGAGAAAGAGAGAGTGTCAAGAGTTTCTTCCAGGAGTGG 744
Db 125 GTTAGAGATGACGAGAAAGAGAGAGTGTCAAGAGTTTCTTCCAGGAGTGG 184
Oy 745 CCAAGCTCCCAAGGCTCCGATGACGAGATCAATGTCGCCACAGAAAGGCTGTTC 804
Db 185 CCAAGCTCCCAAGGCTCCGATGACGAGATCAATGTCGCCACAGAAAGGCTGTTC 244
Oy 805 ATTGAAGTTTGATGATGATCTGTGCTCTCAACAATGACACAAAGCTCTGCAAG 864
Db 245 ATTGAAGTTTGATGATGATCTGTGCTCTCAACAATGACACAAAGCTCTGCAAG 304
Oy 865 TGGGCTGAGAGACAGCTTCCTGATCACTATGCAAGTCTGTGAGAGAGTCTGTCT 924
Db 305 TGGGCTGAGAGACAGCTTCCTGATCACTATGCAAGTCTGTGAGAGAGTCTGTCT 364
Oy 925 CCTGATGCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 984
Db 365 CCTGATGCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 424
Oy 985 GTGCTGTCTCCCGTTACTGTTAGTGAAGAAATCTCCGGGGAACAAAGATCACTTG 1044
Db 425 GTGCTGTCTCCCGTTACTGTTAGTGAAGAAATCTCCGGGGAACAAAGATCACTTG 484
Oy 1045 CTCCTTAGGCGGGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1104
Db 485 CTCCTTAGGCGGGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 544
Oy 1105 AACCGTGAGCTGCTGACACAGTGCCAGGTCGCGCGGCTCTCTATCTGAGTGGCC 1164
Db 545 AACCGTGAGCTGCTGACACAGTGCCAGGTCGCGCGGCTCTCTATCTGAGTGGCC 604
Oy 1165 CTGAGCTGACGAGAGTGTGTGGGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1224
```

```
Db 605 CTGACGCTGACGAGAGTGTGTGGGAGAGAGGTCGCCCCCTTCAACCAAGCTCAAGGC 664
Oy 1225 CTGACGCGGCTTTTGTTTTCATCAGCTCAACCCCTGAGGCGGTGCGGCGGTCTC 1284
Db 665 CTGACGCGGCTTTTGTTTTCATCAGCTCAACCCCTGAGGCGGTGCGGCGGTCTC 723
Oy 1285 AATCTGAGAGAAAGATGTTCTGTAAGCGCTTCTGCGTATGCTGTGAGGAGTGTG 1344
Db 724 AATCTGAGAG-AAAGTTTCTGTTT-AGCGCTTCTGCGGT-TGGTGTGGA-GGAGTGTG 779
Oy 1345 AATGGA 1351
Db 780 AATAGA 786

RESULT 3
BC021272
LOCUS
DEFINITION
Homo sapiens cDNA clone IMAGE:4811303, containing frame-shift
errors.
ACCESSION
BC021272
VERSION
BC021272.2 GI:33878145
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3359)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Bueltow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
Datchenko,L., Marsina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Uebdi,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,D.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalins,D.E.,
Schmerich,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
1247932
2 (bases 1 to 3359)
Director MGC Project.
Direct Submission
Submitted (14-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:16204229.
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigrl.nih.gov
Antler,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
```



Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, D., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,  
Tauben, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>  
Series: IRM, Plate: 39 Row: 9 Column: 3  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 19745161  
This clone has the following problem: frame shifted.

## FEATURES

source

location/Qualifiers

1..3359  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4811303"  
/issue\_type="Placenta, chorioarcinoma"  
/clone\_id="NH MGC\_21"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"

## ORIGIN

Query Match 10.5%; Score 422.4; DB 3; Length 3359;  
Best Local Similarity 50.2%; Pred. No. 9.5e-105;  
Matches 1261; Conservative 0; Mismatches 1206; Indels 45; Gaps 7;

520 TTGGCTGGTCTTTGATTCAGACCGGTGGGCTTCGGGCTCGGACGGTCTTGAC 579  
464 TTGGACCCCTTTTGTCTCCCAAGAACTGGGAAACAGCCAGTACAGTATTTCA 523  
580 GGAAGTCAAGAAATGGAAATCGCTCTAGCCAGAGAGTGTCTGTCTGGCGCA 639  
524 GGACCAAGAAATGGAAACGACCTCTGATGAGCTGATGAGCTGTGGCGAC 583  
640 GTGACCTCTACAGGAGATTTCTCTACGCTTCTCTCCCTTGAAGAGATGAC 699  
584 AACAGATCTTGGGATGATGTTCTGTACAGCTTCTTCTGTGCGAGAACTGAG 643  
700 CGAAGAGAGAGAGATGTCAAGATTCATCTCCAGGAGTGGCCAGACTCCAGGT 759  
644 GAGTGGCCGCAAGAGATTTGGCTGATTTCAAGAGTGGCTGACCCCTGCT 703  
760 CCGGTGACGAGATCATGTCCGACCAAGAGCTGTGTTCATCATTTGACGTTGAT 819  
704 CCAATACAGAGATGTGTCTCAACGAGAGAGACTTGTGTCTATGACAGCTTCA 763  
820 GACCTGGGCTGTCTCAACAT---GACACAAAGCTCTGCAAGAGCTGGGCTGAG 876  
764 GAGCTGACGGGCGCTTGAAAGAACCCGATTCGATCTGTGTGACTGTATGAGAA 823  
877 CAGCTCCGTTCAACCTCATACGAGTCTGTGAGAAAGTCTGCTCCCTGAGCTTC 936  
824 CGGCGGTCAGAGTCTTGTGACGATTTGTCTGAGAAAGAGATCTCCCGAGCTTC 883  
937 CTGATCGTCAACGTCAGAGAGTGGGCAAGAGAGCTCAAGTCAAGAGTGTCTCC 996  
884 CTGCTCATGTCATCAACCGGTGTGCCGAGAGAGCTCCGGGATCAGGTGATCTCA 943  
997 CGTTACTCTTATGATGAGATCTCCGGGAAAGAAATCACTTGTCTTGAAGCC 1056  
944 GAAATCTACAGCCCGGGGATTCACAGAGTGTATGATTTTCTGCTGTTC 1003  
1057 GGGATTTGAGCATCAGAGACACAGAGGTGGTGGTGGATCATGAACACGTCGTC 1116  
1004 TTCAAGAGCCGAAAGAGCATGAGAGCTTCAATCTTTTGAAGAAATGAACAGCTG 1063  
1117 CTGACACAGTCCAGAGTCCCGCGCTGTCTCATCTGCGTGGCCCTCGACGTGAG 1176  
1064 TTTTTCATATGCAATCCCGCTCTCTGTGTGATCTGTGTACAGTCTGAGAGAG 1123  
1177 GACGTGTGGGAGAGGTGCGCCCTTCAACAAAGCTACAGGCTGACGCGCT 1236

1124 ATGCAAGAAAGAAAGACTGAGCTTCCAGACATCTCTGTGTACTCTCT 1183  
1237 TTTGTGTTTCATCAGCTCAACCCCTTGAGAGCGTGTGGCGCTGTCTCATCTGAGAGA 1296  
1184 TTGCTTTTAACTGTTCACACCTGAGGGTGGCCGAGCCGACCTCCGAAACCCAGCA- 1242  
1297 AGAGTGTCTGAGAGCGCTTCTGCGTATGTCTGTGAGAGAGTGTGAATGAGATCA 1356  
1243 -----CGAGCTGAGAGCGCTGTGCTCCCTGCTGAGAGGGTATGTGACAGACATTT 1297  
1357 GGTGTTGATGAGAGAGCTCATAGTTCAGAGACTCGGGAGTGTAGCTCGTCTG 1416  
1298 GAGTTTGTGAAGACGACCTCCGAGAAATGGGTTGTGACCTGACATCTCTGCTG 1357  
1417 TTTCAATGAACATCTTCTCCAGACAGCACTGTGAGAGTACTACATCTTCTCAC 1476  
1358 CTGGGCAACAGATCTCTGAGATGAGGAGCGGTGAGAGCTCTAGTGTCTCCAC 1417  
1477 CTGAGTCTCAGAGATTTCTGTGCGGCTTGTATCTACGTTTGAAGGCTTGAAATCGAG 1536  
1418 GTGTATTCAGAGATCTGTGCGGCTTGTATCTATTTGCTCAAGAGCC-----ACCTTG 1472  
1537 CGAGCTCTGCGCTCTGTAGCTTGAAGAGCAAGAGAGTCCATGAGCTTAAACAGCA 1596  
1473 ATCATCTTACACAGCTGTGAGATGTGACAGAAATCTGATGTCATTT-TCGAAAA 1531  
1597 GGCCTTCATTCATCACTCGCTTGTGATGAGAGGTTCTTGTGTGGCTCGTGTGAGCA 1656  
1532 GCAAGAGAGACATGATGATTTTGTGGGTTTCTTCACTGCGCTTTTAAATAAAG 1591  
1657 GTAGAGAGCCACTGAGAGTCTGTGGCTGTGCTGCTGCTGAGGAGTGAAGAG 1716  
1592 GAACAGAGAAATCTGAGTCCCTTGTGTGCTTCCATCTGCTCAAGATTAAGAGAGCA 1651  
1717 CTCTGCACTGAGTCTCTGTGGGTGAGAGAGCTTAATGCAACACCCAGAGAGAC 1776  
1652 ATTCACAGTCTGAGAGAGCTTGAAGAGGAGTGGCAATCTCAGAGAGAGGTATTC 1711  
1777 CTGAGCGCTTCACTGCTTGTGAGAGCTCAAGCAAGAGATTTGTGCTGTCATTA 1836  
1712 TTGGGATATTTTACTGTCTCTTGAAGATGAGATCTGCTTGTGAGAGAGCAATG 1771  
1837 AACAGCTTCAAGAGTGTGCTTCCATTAACAGAACTGAGACTGTATGATCTTCC 1896  
1772 AACCTCTCAAGAGATCTTCAATTTATTAAGCAAGTGGAGCTGTGTCTCTCC 1831  
1897 TTTGCGCTCAGAGCTGTGCGTATTTGGGAAATTTGGGAGATGTCAGAGAGATCTTC 1956  
1832 TACTGCTTAAATATCTGCTCAAGCTTGAAGAACTGTGTCTTCTTCAAAATGTCTTT 1891  
1957 CCAAGAGATGAGTCCGCTGAGAGCATGTCTGTGCTCTATGAGATCCGGATTAAGCC 2016  
1892 AAGAAAGAGAT-----GAAACACACTCTACGTCCGATTAAGC 1930  
2017 CTGATTGAGAGAGTGGAAAGATTTCTGTCAATGTGACACCCACCACTGCGG 2076  
1931 CTGATCTGT-----TGGCATCATCTGTGCTGTGCTCAACACAGCGGACCTCAGA 1984  
2077 CAGCTGACCTGGGAGAGAGATCTGACAGAGGGGCAAGAGAGCCCTGTGTGGCAG 2136  
1985 GAGCTCCAGGTGACAGAGACACCTTCAGAGTCACTTGTGACCTGTGTAAACAG 2044  
2137 CTGAGGCACTCCCACTGCAAGATACAGACCTGATGTTTGAAGATGACAGATTAACCT 2196  
2045 CTGAGGCACTCCCACTGTGCGCTTCAAGAGCTTGAAGATTAAGTTTCTTTCTGCG 2104  
2197 GGTGTGAGACCTCTGAGAGATCTGTAGGCAACCGTAACTTAAGTCTTCACTTG 2256  
2105 CAGAGTGTCTGCTTGTGAGAGTCTCTTTATAGCCAGACTGAATATCTGAGCTTC 2164  
2257 GAGAGCAACCACTGAGAGAGAGATGATGAGTGTGAGAGCTTAAACACCA 2316

```

Db      2165 ACCCTCAGCAAACTCTCTCGTANGATCAGGTCCTCTGTATGCTTGAACCTGAACCA 2224
Qy      2317 AATGTTGTTGGAGTCTTTGAGGCTGATGCTGTGATTTGACCCATGCTGTATCTG 2376
Db      2225 GCAGGCAACGTCMAAGACTAGCTGCTGTGTAATTTCTACCTCTCTCACTTATGATTGTAA 2284
Qy      2377 AAGATCTCCCAATCTTACGACCTGCCCAAGCTGAAATCTTGAAGCTGGCAGAAAC 2436
Db      2285 GTCCCTGCTGGCTCTTAACCAACAAGAGCTGAGTATCTGAATGTATCTCGCAAC 2344
Qy      2437 AAGGTCAGACACGAGGAGTATGCTCTCAGTATGCTTGAAGTCTCCAGTGGCC 2496
Db      2345 CAGTT--AGACACAGGGGTGCCCTTTTGTGTGAAGCCCTGTGACGCCAGACACGGTTC 2401
Qy      2497 CTGCAAGCTGATATCTGAGGATCTGTGGATCAACAGCAGGGTTCAGAGTCTGSCC 2556
Db      2402 CTGATATCTGATGTTGCTTTCTGCACTCTCAGCAGAGCTGTGGAAATATCTCT 2461
Qy      2557 TCAGCCTCTGACGACCGGAGCTTGAACACCTGTGCTTATCCACAGACCTGGGG 2616
Db      2462 GAAATGCTTCTGCGTAAACAAGAGCGTGGCTATCTAGACCTCAGTCCAAATGCTGAAG 2521
Qy      2617 AACCAAGGTAAATCTATCTGTGTCATCCATGAGGCTTCCCACTGTATGTGAGAGG 2676
Db      2522 GACCAAGGACTGAAAACTCTCTGCGAGGCTTGAACAATCCGAGCTGCTGCTGATTTCA 2581
Qy      2677 CTGATGCTGAATCAAGTGCACCTGTGACACGCGCTGGCTGTGTTCTTCTGCACTTGGCCT 2736
Db      2582 CTGTGTTTGTAAATATGTTTATCATCTGCTGTGCTGTAAGACCTGCGCTCTGCTTC 2641
Qy      2737 ATGGTAACTCATGAGCTGACGCACTGAGCCTTATGATGAACCTGTGAAAGACATGGC 2796
Db      2642 ATCAGCAATCAAAACCTGAATTTCTGCAAAATGGGTGCAATGAATCGAGATGTGGGT 2701
Qy      2797 GTGAAGCTTCTGTGAGAGTATGAGAGAACCATTTTCTCATCTCCAGACCTGAGTGG 2856
Db      2702 GTGAGCTGTGTGTGCGGCTCTGACGATACGATTCGCTTGAATTTCTTGGGTTG 2761
Qy      2857 GTAAAGTCTCATCTTCAACCGCGCGCTGTGAGAGTGTCTGTGTGATCTCGAGAGC 2916
Db      2762 GAAGATGTGGGTTTACGAGACCTGCTGTAAAGATCTGGCTCTGTTCTCACTCGCAGT 2821
Qy      2917 AGACACCTGAAAGAGCTGATCTCAACGCAATGCGCTGGGTGACGATGGGGTGGCTGC 2976
Db      2822 AAGACCTCGACAGAGCTCAACCTGACTTGAACACTTGAACACTTGAACACAGGGGTGGTGTGA 2881
Qy      2977 CTGTGCGAGGAGCTGAACCAAAAGACAGTGTCTTGAAGAGATCTGGGTGA 3028
Db      2882 CTCTGTGAGGCCCTGAGACACCCAGAGTGTGCCCTGACAGTGTCTGCGGCTGA 2933

```

```

RESULT 4
LOCUS   BC012789      3360 bp      mRNA      linear      HTC 12-OCT-2004
DEFINITION
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift
error.
ACCESSION
VERSION BC012789
KEYWORDS
SOURCE   HTC:33872518
ORGANISM Homo sapiens (human)

```

```

REFERENCE
AUTHORS
Strausberg, R.L., Feingold, F.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietzenko, L., Marusina, K., Farmer, A.A., Rubin, A.M., Hong, L.,
Steplikon, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Udell, T.B., Tobin, J.L.,
Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

REMARK
COMMENT

```

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.M.,  
Vallat, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sánchez, A., Whiting, M., Madan, A., Young, J.C., Shcherbko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 3360)  
Director MGC Project.  
Direct Submission  
Submitted (15-AUG-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:15215377.  
Contact: MGC help desk  
Email: [cgabbe-remail.nih.gov](mailto:cgabbe-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadams@systemsbio.org](mailto:amadams@systemsbio.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha  
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

## FEATURES

## SOURCE

```

1..3360
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3957172"
/tissue_type="placenta, choriocarcinoma"
/clone_id="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOT87"

```

## ORIGIN

```

Query Match      10.5%; Score 422.4; DB 3; Length 3360;
Best Local Similarity 50.2%; Pred. No. 9.5e-105;
Matches 1261; Conservative 0; Mismatches 1206; Indels 45; Gaps 7;

Qy      520 TTGCTGTGCTTTTGAATTCGACCGGTGGGCTTCCGGCTGACAGTGTCTGCAC 579
Db      464 TTGACCGGCTTTTGTCTCCAGAGAACTGGGAAACAGCAGCTACATGATATTCA 523
Qy      580 GGAATGTCAGGAATTTGGAATCGGCTCTAGCCAGAAAGATCGTGTGTGGCGCA 639
Db      524 GACACCAAGGAATTTGGAATAACACACACTCCTGATGAAGCTGATGATGCTGTGGAC 583
Qy      640 GTGTGACTTCAACAGGAATGTTCTCTAGCTCTTCTTCTCCCGGTAGAGATGAG 699
Db      584 AACCAAGATCTTTCGGGATGAGTCTCTGATACCGTTATTTCTGCTCAAGAACTAG 643
Qy      700 CGAAGAGAGAGACAGTGTCAACAGTTCATCTCCAGAGAGTGGCCAGACTCCAGGCT 759
Db      644 GAGTTCGCCGCAAGAGATTTGCTGATGATTTCCAGAGAGTGGCCTGACCCGCTGCT 703
Qy      760 CCGGTGACGAGATCATGTCTCCGACAGAAAGGCTGTGTTCATCATATGACGTTTCAT 819

```

D	704	CCTATTAACAGAGATCGTGTCTCAACCGAGAGACATCTTGTGTATCGACAGCTTTCGAA	763
O	820	GACCTGGGCTCTGTCTCAACAT---GACCAAAAGCTTCGAAAGACTGGCTGAAG	876
D	764	GAGCTGACAGGGGGCTTGAAAGAACCCGATTTGGATCTGTGTGTGTAATTGATGAGAA	823
O	877	CAGCTCCGTTCAACCTCATACGAGTCTGTGAGGAAGTCTGTCTCCTGAATCTTC	936
D	824	CGGCGGTGCAGGTCTTCTGAGCAGTTTGTGAGAGAAAGATGTCTCCGAGGCTTC	883
O	937	CTGATCGTCAACCGTAGAGACGTGGGACACAGAAAGCTCAAGTCAAGAGGTCTGTCTCC	996
D	864	CTGCTCATCGGCATTAACCCGTGTGTCCGAGAGAGCTCCGGATCATGGTATGACATCTCA	943
O	997	CGTTACCTGTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTGTCTCTTAGAGCC	1056
D	944	GAAATCTACAGCCCCGGGGATTCAACGAGATGATAGTTAGTGTATTTCTGTGTTTC	1000
O	1057	GGGATTGGTAGCATCAGAGACACAGAGGTTCGTGCGATCATGAAACCGTAGCTG	1111
D	1004	TTCAAAAGACCCGAAAAAGACCATGAGAGCTTCAATCTTTGAAGAAAGATGAAACAGCTG	1067
O	1117	CTGCACAGTGCAGGTGCCGCGTGGCTCTCATCTGGGTGGCCCTGACGCTGACG	1177
D	1064	TTTTCATATGTCAAAATCCCGCTCTCTGTGTGATCTGTGTATCAAGTCTGAAGCAAGG	1122
O	1177	GACGTGTGGGGGAGAGCGTCCGCCCTTCAACCAAGCTCACAGGCTGACGCCGCT	1233
D	1124	ATGCAGAAAGAAAGAACCTGGCCCTGACCTGCAGAGACATCACTCTGTGTATCTCTT	1188
O	1237	TTTGTGTTTCATCAGCTCACCCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGAGGA	1296
D	1184	TTGCTTTTAACTGTTCACACTGAGGGTCCGAGGGCCGACTCCGCAACCAAGCACT	1244
O	1297	AGAGTGTCTCGAAGCGTCTGTCCGATATGCTGTGAGAGGAGTGTGAATGAGAACTCA	1355
D	1243	----CGAGCTGAAGGCCCTGTGTCTCCCTGTGACAGAGGGATGTGTGACACACATTT	1297
O	1357	GTTGTTTANGGAGGACCTCAATGGTTCAAGGACTCGGGAGCTCTGAGCTCCGTGCTG	1411
D	1298	GAGTTTGTGAAGACACTCCGGAGAAATGGGGTTTGTACGCTGACATCTCTGCGCTG	1355
O	1417	TTTCACATGAACATCTTCTCCAGACAGCCCATGTGAGAGTACTTACCTTCTTCCAC	1477
D	1358	CTGGGCAACCAAGTACTTCTGAAATGACGGGAGCGTGAAGCTCTTAAGTGTCTCCGAC	1411
O	1477	CTCAGTCTCCAGGACTTCTGTGCGCGCTTGTATCAAGTGTTAGAGGCGCTGGAATCGAG	1533
D	1418	GTGTGTATTCAGGAGTTCTGTGCGCGCTTGTCTTATTTGTCTCAAGAGCC----ACCTTG	1477
O	1537	CCAGCTCTCTGCGCTCTGTAGCGTTGAGAAGCAAAAGAGTCCATGAGCTTAAACAGGCA	1599
D	1473	ATCATCTTCAACCAAGCTGTGAGATGTATCAGAGAAATGTCTAATTCACATTTTGAAGAA	1533
O	1597	GAGTTTCAATATCCACTCGCTTGTGATGAAGCGTTTCTGTTTGGCTCGTGAAGCAAGAC	1655
D	1532	GCAAGGAGAGACATTTGAGATTTTTTGGGGGTCTTTCTAACGTGGCTTTTAAATTAAGAA	1599
O	1657	GTTAAGAGGCCCATGTGAGGTCTGTGCGCTGTCCCGTTCCCTGGGGGTGAAGCAAGG	1711
D	1592	GAAACAGAAAACTGAGATGCGTTTGTGGCTTCAACGTGTCCCAAGAGATTAAGACAGCA	1655
O	1717	CTTTCGCACTGGGTCTCTGTGTGGGTACAGAGCCTAATGCCACACACCCAGAGAACCC	1777
D	1652	ATTCAACAGTGTCTAAGAGCTTAAGGGAGCGTGGCAATCTCAAGGACAGGTGAATTC	1711
O	1777	CTGACGCGCTTCACTGTCTTTCAGACTCAAGACAAAGATTTGTCGCTTGCATTATTA	1833
D	1712	TTGGGATATTTTATCTGTCTTTGAATGAGAGATCTCGCTTTGTGAAGACGACGAG	1777
O	1837	AACAGCTTCAAGAAAGTGTGCTTCCGATTAACCAAGACTTGAATGACATCTTCC	1899
D	1772	AACCTCTCCCAAGAGCTTAACTTTATTTATTTGCAACGTGACCTGGATGTGTCTTCCG	1833

Oy	1897	TTTCGCCCTCCAGCAGCTGCGCGTATTTTGGGAAAATTTGGGTTGAAATGTCAAAGGATCTTC	1956
Db	1832	TACTGCTTAAATATCTGCTCCAGCTTGAGGAAACTGTGTTTTCCGTCCAANAATGCTTT	1891
Oy	1957	CCAAGATGATGATCCGCTGAGGCAATGTCCTGTGGTCCCTCTATGGATCGGGATTAAGAC	2016
Db	1892	AAGAAAGAGAT-----GAAACAACCTCTACGTCGGAATTACAGC	1933
Oy	2017	CTCATTTGAGAGCAGTGGGAAAGATTTCTGCTCCATGCTTTGGCACCACCAACACTGCGG	2076
Db	1931	CTCATCTGT-----TGGCATCAATCTGCTCTGAGCTCAACCAACGAGGGACACTCAGA	1984
Oy	2077	CAGCTGGAACCTGGGCAAGACATCTCTGACAGAGCGGGCATGMAAGCCTGTGTGCCAAG	2133
Db	1985	GAGCTCCAGGTGACAGACAGACCCCTCAGCCAGTCGACCTTTTGACCTGTGTGTAACAG	2044
Oy	2137	CTGAGGCATCCCACTGCAAGATTCAGACCCGTGATGTTTAAATATGACAGATTAACCCCT	2196
Db	2045	CTGAGGCATCCCACTGCTGCGCTTCAGAGCTTGGATTAATATACGTTTCTTTCTGGC	2104
Oy	2197	GATGTGACAGCACTCTGGAGAAATGTCATGCGCAACCGTAACCTAAGATCCCTCAACTTG	2255
Db	2105	CAGAGTGTCTGCTCTTTGAGGTGCTCTTTTATCAGCCAGACTTGAATTACTGAGCTTC	2166
Oy	2257	GAAGGCAACCACTGAAGAAAGAGATGTAAGATGGCGTGAAGCCTTAAACACCCA	2318
Db	2165	ACCCTCAAGAAACTCTCTCGTGAATGACATCAGAGTCCCTCTGTAGTGCCTTGAATCAACCA	2224
Oy	2317	AAATGTTTGTGAGTCTTTGAGGCTGATGTCGTGATGACCAATGCTGTTAAGT	2376
Db	2225	GCAGGCAACGTCAAGAGACTGACCGCTGTAATTTGTCAACCTTCAACCATTAATTTGAA	2284
Oy	2377	AAGATCTCCCAATTCCTTACAGACTCCCGCCAGCCTGAAATCTCTGAGCCTGGCAGAAAC	2436
Db	2285	GTCCTTGCTGGCCTTCTACCAACAAACAGAGCTGACATCTGTAATGTATCTCGAAC	2344
Oy	2497	CTGCAAGAGCTGATATCTGAGGAACTGTGGCATCAAGCCACGGGTTGCCAAGTCTGGCC	2556
Db	2402	CTGGTATATCTGAAATGTGGCTTTCTTGCCACCTCAGCAGAGCAATGCTCGGAATACATCTCT	2461
Oy	2557	TCAGCCCTCGACACCAACCGGAGCTTGACACACGCTGSCATATCAACAACAGCGGGG	2616
Db	2462	GAAATGCTTCGCGTAAACMAAGAGGTGGCTATCTAAGCTCACTAGTCCAAATGTCCTGAG	2521
Oy	2617	AACGAAGGTGAAATCTACTGTGTGCATTCATGAGGCTTCCCACTGATCTGCAGAGG	2676
Db	2552	GACGAAAGACTGAAAACTCTCTGGAGGCTTGAACATCCGAGCTGTGCTGCTGATTCA	2581
Oy	2677	CTGATGCTGAATCAGTGCACCTGGAACAACGCTGGCTGTGTGTTTTCTTGACCTTGCGTT	2736
Db	2582	CTGTGTTTGGAAATGTTTATCATCTGCTGCTGGCTGTGAAGAATCGCTGCTGCTTC	2641
Oy	2737	ATGGGTAACTATGCTGCTGACGCACTTGACCTTGACATGAACCTGTGGAAGAACAATGSC	2796
Db	2642	ATTCGCAATCAAAACTGMAATTTCTCAAAATTTGGGTGACATGAATTCGAGATGTGGGT	2701
Oy	2797	GTAAGGCTTCGTGACGAGTCAATAGAGAACATCTGTGATCTCCAGCACTGTGAGTGG	2856
Db	2702	GTCAGGCTGTTGTGTGCGGCTCTTAACGCAATTCGGAATTTGCCCTTAAAGATCTTTGGTTG	2761
Oy	2857	GTAAGTGTCACTCACCGCCGCGTGTGTGAGAGTCTGTCTGTGTGATCTCAGAGAGC	2916
Db	2762	GAAAGATGTGGGTTAAAGAGCACTGTGCTGTAAGATCTCGGCTCTGTTCTCACTCGCAGT	2821
Oy	2917	AGACACTGAAGAGCCTGATCTTACGGAACAATCCCTGGGTGAACGTTGGGTTGTCTGCG	2976
Db	2832	AAGACCTGTGACGACTTCACTGACCTTGAACCTTTGACCAACAAGGGGTGTGTGA	2881

Qy 2977 CTGTGCGAGGAGCTGAGCAAAAGACAGTGTCTGACGAGCTCGGGTTGA 3028  
Db 2882 CTCTGTGAGGCGCTGTGAGACCCAGAGTGTGCTCGACGATGCTCGGGCTGA 2933

RESULT 5  
CB228956 814 bp mRNA linear EST 10-FEB-2003  
LOCUS AGENCOURT 11501187 NICHD Rh Ov1 Macaca mulatta cDNA clone  
DEFINITION IMAGE:6884760 5', mRNA sequence.  
ACCESSION CB228956  
VERSION CB228956.1 GI:28280534  
KEYWORDS EST.  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
Cercopithecinae; Macaca.

REFERENCE  
AUTHORS 1 (bases 1 to 814)  
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Elliot Spindel  
cDNA Library Preparation: CLONTECH  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: ILCM3135 row: b column: 23  
High quality sequence stop: 235.  
Location/Qualifiers  
1. .814  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/db\_xref="taxon:9544"  
/clone="IMAGE:6884760"  
/tissue\_type="Ovary"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHD Rh Ov1"  
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;  
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.0-4.0 kb. Tissue pooled from  
pre-pubertal, post pubertal sn menopausal monkeys.  
Constructed by Clontech. Note: this is a NICHD Library."

ORIGIN  
Query Match 10.0%; Score 402.8; DB 6; Length 814;  
Best Local Similarity 88.6%; Pred. No. 1.5e-99;  
Matches 459; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

Qy 2152 TGCAAGATACAGACCTGATGTTTGAATGACAGATTAACCTCGTGTGACGACCTC 2211  
Db 3 TGCAAGATACAGACCTGATGTTTGAATGACAGATTAACCTCGTGTGACGACCTC 62

Qy 2212 TGAGATGTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGCAACCACTG 2271  
Db 63 TGAGATGTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGCAACCACTG 122

Qy 2272 AAGGAGAGGATGTATAGATGCGGTGAAGCTTAACCAACCCCAAAATGTTTGTGAG 2331  
Db 123 AAGGAGAGGATGTATAGATGCGGTGAAGCTTAACCAACCCCAAAATGTTTGTGAG 182

Qy 2332 TCTTTGAGGCTGATGCTGTGATGACCCATGCTCTTAACCTGAAGATCTCCCAATC 2391  
Db 183 TCTTTGAGGCTGATGCTGTGATGACCCATGCTCTTAACCTGAAGATCTCCCAATC 242

Qy 2392 CTTAGACCTCCCGGCTGAATCTCTGAGCTGCGAGGAACAGGTGACAGACAG 2451  
Db 243 CTTAGACCTCCCGGCTGAATCTCTGAGCTGCGAGGAACAGGTGACAGACAG 302

Qy 2452 GAGTAATGCTCTCACTGATGCTTGTAGAGTCTCCAGTGCGCCCTGCAGAACTGATA 2511  
Db 303 GAGTAATGCTCTCACTGATGCTTGTAGAGTCTCCAGTGCGCCCTGCAGAACTGATA 362

Qy 2512 CTGG-AGACTGTGGCATCACAGCCAGCGTTGGCAGAGTCTGGCCCTCTGTCAG 2570  
Db 363 CTGGAGAGACTGTGGCATCACAGCGGTTGGCAAAAGTCTGGCCCAACCTCTGTCAG 422

Qy 2571 CAACCGAGGCTTGACACACCGTGTGCTATCCAAACA-CAGCTGTGGGAGCAAGGTGAA 2629  
Db 423 CTACCGAGGCTTGACACACCGTGTGCTATCCAAACAAGGAGAGAGAGTGA 482

Qy 2630 ATCTACTGTGTCATCCATGAGGCTTCCCACTGACT 2667  
Db 483 AACCGCTGTATCATCATGAGGTTTCCCATGACT 520

RESULT 6  
COB14088 776 bp mRNA linear EST 06-AUG-2004  
LOCUS AGENCOURT 30247088 NIH\_MGC\_256 Mus musculus cDNA clone  
DEFINITION IMAGE:30936082 5', mRNA sequence.  
ACCESSION COB14088  
VERSION COB14088.1 GI:51032714  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS 1 (bases 1 to 776)  
TITLE NIH-MGC http://mgc.nci.nih.gov/.  
TUMOR National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: NDAM1191 row: d column: 11  
High quality sequence stop: 724.  
Location/Qualifiers  
1. .776  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30936082"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH MGC 256"  
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;  
Site 2: NotI; cDNA was primed using oligo-dt primer:  
5'-pGACTGATGTTAGATTCGAGCGGCGCCGCTT25-3' and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb  
resulted in an average insert size of 1.2 kb. This is a  
primary library (normalized primary library is NIH MGC 257)  
and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

ORIGIN  
Query Match 9.5%; Score 381.4; DB 7; Length 776;  
Best Local Similarity 70.3%; Pred. No. 1.3e-93;  
Matches 511; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

Qy 2468 GTGATGCTTTAGAGTCTCCAGTGCCTGCGAGAAAGCTATATCTGAGAGACTGTGCA 2527  
Db 3 GGTATGCTTTAGAGTCTCCAGTGCCTGCGAGAAAGCTATATCTGAGAGACTGTGCA 62

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title
2528	1	mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493411p06 product:maternal effect gene, full insert sequence.	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2588	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2648	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2708	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2768	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2828	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2888	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
2948	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3008	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3068	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3128	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3188	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3248	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3308	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3368	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3428	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3488	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3548	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3608	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3668	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3728	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3788	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3848	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3908	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
3968	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4028	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4088	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4148	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4208	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4268	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4328	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4388	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4448	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4508	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4568	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4628	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4688	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4748	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4808	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	AK016782
4868	1	mus musculus (house mouse)	AK016782	1913 bp mRNA	linear	HTC 03-APR-2004	

JOURNAL MEDLINE PUBMED	prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE AUTHORS	3 Shihara,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komno,H., Akiyama,J., Nihi,K., Kitanishi,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kawabagi,K., Fujikake,S., Inoue,K., Togaw,Y., Izawa,M., Ohara,E., Watabiki,M., Yonezaki,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsumura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kitara,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL MEDLINE PUBMED	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1913)
TITLE	Aadachi,T., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arikawa,T., Bono,H., Carninci,P., Fukuda,S., Fukushima,Y., Funano,M., Hanigaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hirakawa,T., Horii,F., Imocani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukabe,T., Kato,H., Kawai,J., Kojima,Y., Komno,H., Kouda,M., Koya,S., Kuuhara,C., Matsuyama,T., Miyazaki,Y., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,K., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE JOURNAL	Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site ( <a href="http://genome.gsc.riken.jp/">http://genome.gsc.riken.jp/</a> ) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAAGGAATCGATCAAGACCTCTTTTCTTTTTTNN 3']. cDNA was prepared by using trehalose chemo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAAGCATTCGCAGTTAAATTAAATTAATCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size compressed longer than 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from lambdaB FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B. Location/Qualifiers 1. 1913 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:4933411P06" /db_xref="taxon:I0090" /clone="4933411P06"
FEATURES SOURCE	

```

/sex="male"
/cisue_type="testis"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
735..1544
/note="unnamed protein product; maternal effect gene
(MGP MGI:1345193)
putative"
/codon_start=1
/protein_id="BAB30427.1"
/db_xref="GI:12855703"
/translation="MKDDMKLACEALRHPKCALRNLDSGLTPASCHLVSALES
NONTPLCLSNNSLSTEGVQOLCOPIRPECALRLINHCNIDMDNAGFALALANN
TKLTHLSITMNPVPGAMKLCALKEPTCKLKEKLVDCOLMOCNCCGLACMTITTK
HLKSLIDGNNAIDKGVITLCEGLKROSSSRIRLIGLACETLSNCBSLSAISCNPH
LNSLNLKNDPSTSGKRLKCSAFQCPVSNLGIISGNLSFALFIALRRR"
polya_signal
1886..1891
/note="putative"
1913
/note="putative"
ORIGIN
Query Match          9.4%; Score 377.8; DB 3; Length 1913;
Best Local Similarity 65.7%; Pred. No. 1.8e-92;
Matches 550; Conservative 0; Mismatches 287; Indels 0; Gaps 0;
QY 2381 TCTCCAAATCTTAACGACCTCCCGACCTGAATCTCTGAGCCTGGCAGAAACAAG 2440
DB 676 TCTGGAACCTCTTTTGAACATCAAACTTAAGTACTCAATTAAGGAAACCCCA 735
QY 2441 TGACAGACGAGGAGTAATGCTCTAGTGAATGCTTGAAGATCCAGTGCCTGCG 2500
DB 736 TGAAGAGTGAACATGAATTAAGCTCGAAGCGCTGAACATCAAGTCCGCCCTCG 795
QY 2501 AGAAGCTGATCTGAGAGCTGTGSCATCAACGCCGTTGCCAGATCTGCGCTGAG 2560
DB 796 AGGCTCTGAGACTGACAGCTGTGCGCTCAACCTGCAGCTGCACCTTCTGCTCAG 855
QY 2561 CCTCTGTCAGAACCGGAGCTTGAACACCTGTGCTTCAACAACAGCCTGGGGAACG 2620
DB 856 CCTTTTGAAGAACGAACTTGACACACTGTGCTTCAACAACAGCCTGGGGAACG 915
QY 2621 AAGGTGAATCTAATCTGTGTCATGATGAGGCTTCCCACTGATCTGACAGGCTGA 2680
DB 916 AAGAGTGCAACAGCTGTGTCACTTCTGAGGAATCCAGATGTCTCTCAACGGCTGA 975
QY 2681 TGCTGAATCAAGTCACTGACACAGCGCTGCTGTGTTTCTTGACCTTGCGCTTATGG 2740
DB 976 TACTGAATCAATGCAACATTTGATGATGATGCTTATGCTTCTGCACTTGAGACTTGA 1035
QY 2741 GTAATCTATGCTGACGACCTGAGACCTTGAACATCCCTGTGGAACAATGCGCTGA 2800
DB 1036 ACAACAAAGGTGACCACTGACCTGACCATTAACCCGTAGGGAGTGTGAATGA 1095
QY 2801 AGCTTCTGTCAGAGCTCATGAGAACCATCTTGCATCTCCAGAGACTGAGTGTGATA 2860
DB 1096 AGCTACTGTGTGAAGCTTTAAAGAACTTACTTGTACTTAAGAACTGAACCTAGTGG 1155
QY 2861 AGTGTCAATCAACCGCGCTGTCGAGAGTCTGTCTGTGTGATCTCGAGAGACAGAC 2920
DB 1156 ACTGCAACTTATGAGAACTGCTGTGAGACCTGGCTGTATGATCAACAACCAAGC 1215
QY 2921 ACCTGAAGAGCTGATCTCAACGACATGCTTGGGTGACGAGTGGGTGTGTGGCTGT 2980
DB 1216 ACTTAAAAAGTTTGGATCTTGGTAAACAAGCCCTGGTGAACAAGAGTCAATTAACCTGT 1275
QY 2981 GCGAGGAGCTGAAGCAAAAGAACAGTGTTCGACGAGACTGCGGTTGAAGGAGATGAGAC 3040
DB 1276 GCGAGGAGCTGAAGCAAAAGTGAAGCACTCCCTGAAGAGACTTGGGTTGGGGGATGTAGT 1335
QY 3041 TGACTTCTGATTTGCTGTGAGGCACTCTCTTGGCCCTTTCCTGCAACGCGCATCTGACCA 3100
DB 1336 TGACTTCAATTTGCTGTGAGTCAATGTATGGCCATCTTGGCAACCTGACCTGAACA 1395

```

```

QY 3101 GTCTAACTGTGTCAGAAATTAATCTAGTCCCAAGGAATGATGAGCTGTGCGGCT 3160
DB 1396 GCTTAACCTGATATAAGAAATGACTTCAGTACATGGGGAACTTAAGCTGTGCTGCT 1455
QY 3161 TTGCTGTCCCACTCTTAATCAAGATTAATTTGGCTGTGGAATGGCAGTACCTTG 3217
DB 1456 TCGAATGCCCTGTCTTAACCTGGGAGATTAATTTGGGTGTGGCACTTGAGCTCACCTG 1512

RESULT 8
CO804754 808 bp mRNA linear EST 05-AUG-2004
LOCUS AGENCOURT 30257031 NIH MGC 256 Mus musculus cDNA clone
DEFINITION IMAGE:30938137 5', mRNA sequence.
ACCESSION CO804754
VERSION CO804754.1 GI:50992934
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 808)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga9bs-remail.nih.gov
Tissue Procurement: Dr. Kathleen Horner, Stanford University
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM196 row: j column: 02
High quality sequence stop: 742.
Location/Qualifiers
1..808
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30938137"
/lab_host="DH10B Tona"
/clone_id="NIH_MGC_256"
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;
Site 2: NotI; cDNA was primed using oligo-dT primer;
5'-pGACTAGTTCAGATGCGAGCGGCCGCCCT(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary library (normalized primary library is NIH MGC 257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"
ORIGIN
Query Match          9.3%; Score 374.4; DB 7; Length 808;
Best Local Similarity 70.0%; Pred. No. 1.1e-91;
Matches 504; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
QY 2462 CTCACAGTATGCTTGAAGATCTCCAGTGCAGCGCTTCGACAACTGATTAATCTGAAGACT 2521
DB 10 CCTTGGGAATGCCCTTGAAGTCAATGTCTTACTGCAAAAGTTGATCTGACAACT 69
QY 2522 GTGGCATCAGACGACGAGGTTCAGAGTCTGAGCTGAGCCCTGTCGACCAACGAGACT 2581
DB 70 GTGGCTACACCTGCGCAAGCTGCAACCTTCTGAGTCTAGGCCCTTTTCAAGCAACGAAGT 129
QY 2582 TGACACACTGTGCTATTCACAACAGCCTGGGGAAGGAAGTGAATTAATCTACTGTGTC 2641
DB 130 TGACACACTGTGCTGCTGCAACAGCCTGGGGAAGTGAAGTGAACAGCTGTGTC 189

```



OY	2642	GATCCATAGAGGCTTCCCACTGTA	CTCGCAGAGGCTGATGTAATCAGGCCACTGG	2700		
Dp	190	AGTTCCTAAGGAATCAAGATGTGCT	CTCCAGCGGCTGATTAATCACTGCACAATTG	249		
OY	2702	ACACGGCTGAGCTGTGGT	TTTCTTGCAC	TTCGCGCTTATGGGTAACTCATGCTGACGCACC	2765	
Dp	250	TAGATGATGCTTATGTGCTT	CCCTGGCAATGAGACTTGGCAACACCAAGCTGACCCACC	309		
OY	2762	TGAGCTTTAGCATGAA	CCCTGTGGAAAGCAATGCGTGAAGCTTGTGTCGAGGTCATGA	2821		
Dp	310	TGAGCTGACCATGAA	CCCCGATAGGGAGATGGTGAATGAAAGCTACTGTGTGAAGCTTTAA	369		
OY	2822	GAGAACCACTTTGTATCTTCCAGGA	CCCTGGAAGTTGTAAAGTGTATCTCACCGCCGGT	2883		
Dp	370	AGGAACCTACTGTGTAACTTCA	AGGAACCTGGAACCTGAGAACTGTAGCACTGCCAACCTCACAGAACT	429		
OY	2882	GCTGGAAGATCTGTCTGTGTGATCT	CGAGGAGACA	CACTGGAAGGCTGTGATCTCA	2941	
Dp	430	GCTGGAGAGACCTGCGCTGTATGATCA	CAACACCAAGCACTTAAAAAGTTGGATCTTG	489		
OY	2942	CGGACAATGACCCTGGGTGACCGGTGGGGTGTGCTG	CGCTGTGCGAGGGA	CTGAAGCAAAAGA	3001	
Dp	490	GTAACAAAGCCCTGGGTGACAAAGGAGTCA	TAACTCTGTGTAGGGA	CTGAAGCAAAAGTA	549	
OY	3002	ACAGTGTCTTGA	CGAGACTCGGGTTGAAGCA	TGTGA	CTGACTTCTGATTTGCTGTGAGG	3061
Dp	550	GCAGCTCCCTAGAGAGACTTGGGGT	TGGGGGGCATGTAA	GTGACTTCCAAATTTGCTGTGAGG	609	
OY	3062	CAC	TCTCTTGGCCCTTTCTGTGCAACGGGCATCTGA	CAAGTCTAAACTGGTGCAGAA	TAA	3121
Dp	610	CATTGTCA	TGTGGCCATCTCTTGCAACCTTCA	CTTGAACAGGCTTAAACCTGGTGAAGAA	TG	669
OY	3122	ACTTCAGGCCA	AAAGGATGATGAAAGCTGTGTGCGCCCTTGGCTGTGCCACTTA	CT	3181	
Dp	670	ACTTCAGATCA	TCGGGGATGTTGAAGCTGTGCTGTGCGCTTCCAAAGCCCTGTCTTA	CT	729	

RESULT 9	BU634350/c	LOCUS	DEFINITION
BU634350	693 bp	mRNA	linear
UI-H-FL1-bgx-m-16-0-UI.s1	NCI	CGAP FL1 Homo sapiens	CDNA clone
UI-H-FL1-bgx-m-16-0-UI 3'			mRNA sequence.

ACCESSION	B0634350	GI:23301605
VERSION	B0634350.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 693)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.

Email: [csapbs-remail.nih.gov](mailto:csapbs-remail.nih.gov)  
Tissue Procurement: James Martin  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
POLYA=yes.

```
FEATURES
source
location/Qualifiers
1. .693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-Fli_bgx-m-16-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
```

```

/!lab host="VDH10B (Life Technologies)"
/!clone lib="NCI CGAP FL1"
/!note="Organ: Chondrosarcoma; Vector: pRTT3-Pac
(Pharmacia) with a modified polylinker; Site_1: EcoR I;
Site_2: Not I; NCI CGAP FL1 is a normalized cDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GAGCTCGCTG. The cell lines were provided by Dr. James
Marrin from the University of Iowa.
TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_Lib=ur-H-FL1
TAG_SEQ=GAGCTCGCTG"

```

Query Match	8.9%	DB 5	Length	693;
Best Local Similarity	98.7%	Pred. No.	3.3e-87;	
Matches	371;	Conservative	0;	Mismatches 4;
				Indels 1;
				Gaps 1

293 / TCTCAGCGACAATGCCCTGGGTGACGGTGGGGTGTCTGCGCTGTGCGAGGGACTGAAGCA 2996

0931C1C1CAGGACAA1GCCCC1GCGG1GACAG1GGGG-1G1C1GAC1G1GCGAGAGAC1GAGAGCA 635

[illegible]

3057 TGAGGCACTCTCCTTGCGCCCTTTCCTGCAACCGGCACTGACCAGTCTAACCTGGTGCA 3116

Db 574 TGAGGACCTCTCTTGCCCTTCTGCAACGGCATCTGACCACTTAACCTGGTGCA 515

3117 GAATTACTTCAGTCCCAAGGAATGATGAGCGTGTTCGGCCYTTGCTGTCCACGTC 3170

[illegible]

454 TAACTTACAGATAATTCGCTGTGAGAAATTCAGTACCCTGTGCAAAATTAAGGAAAGCTGCT 395

3237 GGAGGAGTCAGCTACTCAAGCCCGAGTCGTAATTGACGCTAGTTGGCATTCTTTGA 3296

Db 394 GGAGGAAGTGCAGCTACTCAAGCCCGAGTGTAAATTGACGGTAGTTGGCATTCTTTTGA 335

329/ TGAGATGACCGACAC 3312  
|||||

RESULT 10	BU630481/c	LOCUS	DEFINITION
BU630481	684 bp	mRNA	linear
UI-H-FL0-bdk-c-22-0-UI	81	NCI_CGAP_FLO	Homo sapiens
UI-H-FL0-bdk-c-22-0-UI	3		mRNA sequence.
			EST 23-SEP-2003

ACCESSION	BU630481	GI:23297486
VERSION	BU630481.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 684)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: James Martin  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@iowa.edu  
Seq primer: M13 FORWARD  
PolyA=Yes.

## FEATURES

source

Location/Qualifiers  
1..684  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UT-H-FL0-bdk-c-22-0-UT"  
/tissue\_type="Cell lines"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP FL0"  
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac  
(Pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI\_CGAP\_FL0 is a cDNA library derived from  
a pool of mRNA obtained from 4 cell lines from grade III  
chondrosarcoma tissues. The library was constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT7T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GAGTCGGTG. The cell line  
was provided by Dr James Martin from University of Iowa.  
TAG\_TISSUE=Human Chondrosarcoma Grade 3 cell line mix  
TAG\_Lib=UI-H-FL0  
TAG\_SEQ=GAGTCGGTG"

## ORIGIN

Query Match 8.9%; Score 357.2; DB 5; Length 684;  
Best Local Similarity 98.9%; Pred. No. 6.1e-87;  
Matches 370; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 2999 TCACGGAATGCGCTGGGTGACGGTGGGCTTGCCTGCGAGGACTGAAGCAAA 2998  
DB 684 TCACGGAATGCGCTGGGTGACGGTGGGCTTGCCTGCGAGGACTGAAGCAAA 626  
QY 2999 AGAAGAGTCTCTGACGAGACTCGGTTGAAGGACATGGACTGACTTCTGATTGCTGTG 3058  
DB 625 AGAAGAGTCTCTGACGAGACTCGGTTGAAGGACATGGACTGACTTCTGATTGCTGTG 566  
QY 3059 AGGCACTCTCTGGCCCTTTCCTGCAACCGGCATCTGACAGTCTAAACCTGGTGCA 3118  
DB 565 AGGCACTCTCTGGCCCTTTCCTGCAACCGGCATCTGACAGTCTAAACCTGGTGCA 506  
QY 3119 ATTAATCTAGTCCCAAGGAATGATGAAGCTGTGTCGGCCCTTTCCTGCAAGCTCA 3178  
DB 505 ATTAATCTAGTCCCAAGGAATGATGAAGCTGTGTCGGCCCTTTCCTGCAAGCTCA 446  
QY 3179 ACTTACAGATATTGGGCTGTGGAATGAGCACTCTGTGCAAAATGAAGAGCTGTG 3238  
DB 445 ACTTACAGATATTGGGCTGTGGAATGAGCACTCTGTGCAAAATGAAGAGCTGTG 386  
QY 3239 AGGAAGTGAAGTCACTCAAGCCCGAGTGTGTAATGACGGTAGTGGCAATCTTTTGATG 3298  
DB 385 AGGAAGTGAAGTCACTCAAGCCCGAGTGTGTAATGACGGTAGTGGCAATCTTTTGATG 326  
QY 3299 AAGATGACCGACAC 3312  
DB 325 AAGATGACCGGTAC 312

## RESULT 11

CO802335

LOCUS

## DEFINITION

CO802335

ACCESSION

## VERSION

CO802335.1

## KEYWORDS

EST.

## SOURCE

Mus musculus

## ORGANISM

Mus musculus (house mouse)

## REFERENCE

1 (bases 1 to 786)

## AUTHORS

NIH-MGC

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDAMI201 row: k column: 11  
High quality sequence stop: 505.

## FEATURES

source

Location/Qualifiers  
1..786  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30940090"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH MGC 256"  
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;  
Site 2: NotI; cDNA was primed using oligo-dT primer;  
5'-pGATGTTCTTCAATCGGACGGCCGCTT25-3 and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb  
resulted in an average insert size of 1.2 kb. This is a  
primary library (normalized primary library is NIH MGC 257)  
and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 8.7%; Score 349.2; DB 7; Length 786;  
Best Local Similarity 69.3%; Pred. No. 1.1e-84;  
Matches 474; Conservative 0; Mismatches 210; Indels 0; Gaps 0;  
QY 2462 CTCCTGATGATGCTCTGAGAGTCTCCAGTSCGCCCTGCGAGACTGATATCTGAGAGACT 2521  
DB 1 CCTTGGGAATGCCCTTGTGAGTCAATGTCTACTGCAAAAGTGTACTGCAACA 60  
QY 2522 GTGGATCAGCAGCAAGCGGTTGGCAGAGTCTGAGCCCTGCTGAGCAACCGGAGCT 2581  
DB 61 GTGGATCAGCAGCAGCGGTTGGCAGAGTCTGAGCCCTGCTGAGCAACCGGAGCT 120  
QY 2582 TGAACACCTGTGCTATCCAAACAAGCCTGGGGAAGGAAGTGAATCTACTGTGTC 2641  
DB 121 TGAACACCTGTGCTATCCAAACAAGCCTGGGGAAGGAAGTGAATCTACTGTGTC 180  
QY 2642 GATCCATGAGGCTTCCCACTGTAGTGTGACAGAGGCTGTGATCAATGAGTCACTG 2701  
DB 181 AGTTCTGAGGAATCCAAATGTCTCTCCAGCGGCTGTGATCAATCACTGCAACATTTG 240  
QY 2702 ACAAGGCTGTGCTGTCTTTTCTTCACTGTGCGCTTATGGGTAATCAATGCTGACGAC 2761  
DB 241 TAGATGATGCTTATGGCTTCTTGGCAATGAGACTTGGCAACAACAAGCTGACCCACC 300  
QY 2762 TGAGCTTATGATGAACCTGTGGAAGCAATGCGTGAAGCTTCTGTGCGAGGTCAATGA 2821

```

Db      301  TAGGCTGACCAATGAACCCCGTAGGGGATGGGCAATGAGTACTGTGAAAGCTTTAA 360
Qy      2822 GAGAACCATCTTGTATCTTCACAGACCTGAGATTGGTAAAGTCACTTCACGCCCGCT 2881
Db      361  AGGAACCTTACTTGTATCTTCACAGACCTGAGATTGGTAAAGTCACTTCACGCCCGCT 420
Qy      2882 GCTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACACCTGAAAGAGCTGTGATCTCA 2941
Db      421  GCTGAGAGAGACCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy      2942 CGAACAATGCTCTGGGTGACGAGTGGGTTGCTGCGCTGTGCGAGAGAGAGAGAGAGAG 3001
Db      481  GTAAACAAGCCCTGGGTGACAAAGAGTCAATACCTGTGTGAGAGAGAGAGAGAGAGAGAG 540
Qy      3002 ACAGCTTCTGACAGACCTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3061
Db      541  GCAGCTCCCTGAGAGACTTGGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy      3062 CACTGCTCTGGGCTCTTCTGCAACCGGCACTGTGACGATCTTAACCTGTGAGAGAGAG 3121
Db      601  CATTGCAATGGCCATCTTTCGACACCTCTCACTGAACAGCTTAACCTGTGAGAGAGAGAG 660
Qy      3122 ACTTCACTCCCAAGAGATGATGA 3145
Db      661  ACTTCACTATCATCGGAGATGTTGA 684

```

```

RESULT 12
LOCUS    CO382844
DEFINITION AGENCOURT 26622956 NIH_MGC_253 Rattus norvegicus cDNA clone
IMAGE:7303813 5', mRNA sequence.
ACCESSION CO382844
VERSION   CO382844.1 GI:49490946
KEYWORDS  EST.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.

```

```

REFERENCE 1 (bases 1 to 792)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Daniela S. Gerhard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10N07 Bethesda, MD 20892
          Email: gsa@bbs-r@mail.nih.gov
          Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
          College of Wisconsin
          cDNA Library Preparation: Express Genomics
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLNL15339 row: 9 column: 11
          High quality sequence stop: 690.

```

```

FEATURES
source
1..792
location/Qualifiers

```

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7303813"
/tissue_type="Ovary - Brown Norway Line 3 Age 8 weeks.
Tissues were snap-frozen and transferred in -70C. RNase
free the entire procedure."
/lab_host="DHI08 Tona"
/clone_1ib="NIH MGC 253"
/note="Organ: ovary; Vector: pExpress-1; Site_1: EcorV;
Site_2: NotI; RNA obtained from female animals at 8 wk

```

old. Tissues were snap-frozen and kept at -80C for two days before RNA extraction and purification (Tri-reagent method). cDNA was primed using oligo-dT primer: 5'-pACTAGTCTTAGATGCGAGCGCGCC(7)25-3' and cloned into the BcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.5 kb. This primary library is normalized (non-normalized primary library is NIH MGC 252) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH\_MGC library."

#### ORIGIN

Query Match 8.3%; Score 336; DB 7; Length 792;  
Best Local Similarity 66.6%; Pred. No. 4.9e-81;  
Matches 480; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```

Qy      2102 TGACAGAGCGGCGCATGAAGACCTGTGTGCCAGCTGAGCATCCACCTGCAGAGATG 2161
Db      2  TGAATGAATGCTCCATGAAGATTTATGCTTCAACCTGAGAAATTCATCTTCACATAC 61
Qy      2162 AGACCTGATGTTTGAATGACAGATTACCCCTGTGTGTGAGAGACCTTGTGAGATG 2221
Db      62  ABAATCTTAACGTTTAAAGATTCAGAGTAGTTTCTGGGCTGCAATATCTGTGAGATGCTCC 121
Qy      2222 TCATGGCCCAACCGTAACCTTAAGATCCCTCAACTGGAGGACCCACCTGAAGAGAG 2281
Db      122  TTGTTAGCAATGGAGAACTTAAGTACTCAATCTTAGGAAACAGCCCATGAAGAGAGATG 181
Qy      2282 ATGTAAGAGATGCGCTGTGAGAGCTTTAAACACCCAAATGTTTGTGAGATGCTTTGAGGC 2341
Db      182  ACATCAAGTTAGCTGTGAGAGCATGAACATCCAGAGCTGTTCCCTGAGAGACCTGAGAGT 241
Qy      2342 TGGATTGCTGTGTGATGAGACCCATGCTGTACTGAATCTCCCAATCCTTAGAGACT 2401
Db      242  TGGATTCTGTGAGATTACCTCACTGTGTATGAGACTATCTCCAGAGCTCTTCTTTAG 301
Qy      2402 CCCCAGGCTGAATATCTGTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 2461
Db      302  CCAGAGGCTTAAGATGCTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGATG 361
Qy      2462 CTCTGATGATGCTCTTGAAGAGTCTCCAGTGGGCTGTGAGAGAGATGATGAGAGACT 2521
Db      362  CCTTGGGAGAGCTTGTGAGTACTTATCATGTAACCTCAAGAGCTGATGAGAGCT 421
Qy      2522 GTGGATCAACAGCCAGGCTTGCAGAGTCTGAGCTTCAAGCTTCTGCAAGACCGGAGCT 2581
Db      422  GTGACCTATATCTGTGAGCTGTGAGAGCTGAGCTTCCGCTCTCCAGAGACCGGAGCT 481
Qy      2582 TGACACACCTGTGCTTATCCAAACAGCGCTGGGAGAGAGAGTGAATCTACTGTGTC 2641
Db      482  TGACTACCTGTGCTTGTCAAAACAGCTTGGGAGCTGAAGAGAGAGAGAGAGAGAGAGAG 541
Qy      2642 GATCCATGAGGCTTCCCACTGATGCTGAGAGAGCTGATGCTGAATGAGTGCACCTGG 2701
Db      542  AGTCTCTGAAGAAACCAAAATGTGCTTCGCGGCTGATATGATGATGATGATGATGATGAT 601
Qy      2702 ACAGCGCTGGCTGTGTGTTTCTGTGACTTGGCTTATGAGTGAATGATGATGATGATGAT 2761
Db      602  TCAAGAGATGCTTACGCGCTTCTGCGCTGATATGAGCAAAACAGAAAGCTGACGACAC 661
Qy      2762 TGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2821
Db      662  TGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
Qy      2822 G 2822
Db      722  G 722

```

```

RESULT 13
LOCUS    BM454123
DEFINITION AGENCOURT 6419582 NCI CGAP_Ov44 Mus musculus cDNA clone
IMAGE:5504420 5', mRNA sequence.

```

ACCESSION BM454123  
VERSION BM454123.1 GI:18503163  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1119)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
TITLE NIH-MGC  
AUTHORS  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Aaron Hsueh  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: LLM12145 row: 1 column: 21  
High quality sequence start: 3  
High quality sequence stop: 705.  
Location/Qualifiers  
1..1119  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5504420"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_1ib="NCI CGAP Ova4"  
/note="Organ: ovary; PMSG-treated; Vector: pCMV-SPORT6.1; Site\_1: EcorV; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.2 kb. Library constructed by Life Technologies. Note: this is a NCI CGAP library."  
ORIGIN  
Query Match 8.0%; Score 323.4; DB 4; Length 1119;  
Best Local Similarity 62.7%; Pred. No. 176-77;  
Matches 605; Conservative 0; Mismatches 346; Indels 14; Gaps 6;  
QY 842 ATGACACAAAGCTCGACAAAGCTGGGCTGAGAGACACCTCCGTTCAACCTCATACCA 901  
DB 13 ATGATATACACTATCCAGAGACTGGAAGATGAACACCCATATACATCCGATGTACA 72  
QY 902 GTCTGTGAGGAAGGCTCTCTCTCTCTGAGTCCTTCTGATGTCAACCGTCAGAGAGTGG 961  
DB 73 GCTCTCTGAGGAAGGCT 132  
QY 962 GCACAGAGAGCTCAAGTCAGAGGTCGTCTCCCGTTACTCTTACTTATAGAGATCT 1021  
DB 133 GCTTAGAAAACCTCAAGTCATAGTGTGTGTCCTCTATATACTGTGAGAGACTGT 192  
QY 1022 CCGGGGAACAAAGATCACTTGTCTTGAAGCGGGGATTTGGATGAGATCAGAGACAC 1081  
DB 193 CTGATCAAGAGATCTCAGCTGTCTCTGAGAACATCTCCATAGAGCTGATAGAAATC 252  
QY 1082 AAGGGTTCGTGATCAATGAACAAACCGTGAAGCTCTGACCAAGTCCAGGTCGCGC 1141  
DB 253 AAGTTCCTCATCTCTGATAGAAATCAACAGCTCTTGAACCAATGCCAGGCCCTCTG 312  
QY 1142 TGGGCTCTCATCTGCGCTGGCGCTGAGCTGACAGAGAGTGGGGAGAGCGTGGCC 1201  
DB 313 TGTGCTCTCTGTGTGAGGCTTACAGCTACAAAGAACTGGGAAAAGATGCACCC 372  
QY 1202 CTTTCAACCAAGCTCAAGGCTGACGCGCTTTTGTGTTTATGATGAGCTCAACCTTC 1261  
DB 373 TACCTGCAAGACTCTACCGGTTTGTATGACACCTTGTGTTTACACGCTCACTTGA 432  
QY 1262 GAGGCGGTGTCGCGGCTGTCTCAATCTGAGAGAAAGTGTCTCTGAAGCGCTTCTGCC 1321  
DB 433 AAAGGCTTCCAGAGCGCTCTCAGTCAAGAAAGACAGATTACTAGTGGGTTGTGCA 492

QY 1322 GTATGCTGTGAGGAGGTGTGAATAGAAAGTCACTGTTTATGATGAGACGACTCATGG 1381  
DB 493 TGATGGCAGCTGAAAGAGGTGTGACATAGAGTGGTGTCTTATGATGATGACTGAAGA 552  
QY 1382 TTCAAGAGCTCGGGAGCTGAGGCTCGTGTCTGTTTCAATGAAACATCTTCTCCAG 1441  
DB 553 ACTATAGGCTTAAAGAGCTGAGATCTTGGCCCTCTTTCATGATGACATCTTCTCCAG 612  
QY 1442 ACAGCC--ACTGTGAGAGTACTACACTTCTTCCACTTCAAGTCTCAAGACTTCTGTG 1498  
DB 613 TTGGCCAAACAGGACAGTGTATGTTTCTCCACCTCAGCTCAGGATTTCTTGG 672  
QY 1499 CCGCTTGTACTAGCTGTGTTAGAGGGCCGTGAATATGACGACAGCTCTGACCTCTGACG 1558  
DB 673 CTGCTTATATTAATGTTTATGAGAGGCTGAG--GGATGATACAGATTTTGTCTTCA 729  
QY 1559 TTGAGAAACAAAGAGGTTCATGAGCTTAAACAGGAGGCTTGCATATCCATCGCTTT 1618  
DB 730 TTGAAAACAAAGAGGATCATGAGGTGA--AGAGAACTGACGACCTGCTTCTCG 786  
QY 1619 GGATGAAGCGTTTCTTGTGTTGGCTCTGAGCGGAAGACGTAAAGAGGCACT--GAGGTC 1677  
DB 787 GGATGAAGCGTTTCTTATTTGGCTCATGAACAAAGATATCTTGAAGACTCTGAGAGGTT 846  
QY 1678 CTGCTGGGCTGTCCCGTTCCCTGCGGGGTGAAGCAGAGCTTGTGACTGGGCTCTCTG 1737  
DB 847 CTGTTTGAATATCCCGTATTCCTGATTCCTGT--TGACAAAGCTCCACACTGGGTCTCTG 904  
QY 1738 TTGGGTACAGACCTTAATGACACCCACGAGAGACACCTGAGCGCTTCTCACTGTCT 1797  
DB 905 ATACTGACGAGGCGCATGAGCACCGCCCAA--TGACACCTGAGATGCTTCAATTGGCTA 962  
QY 1798 TTGCA 1802  
DB 963 TTGCA 967  
RESULT 14  
LOCUS COB15646 739 bp mRNA linear EST 06-AUG-2004  
DEFINITION AGENCOURT\_30254302 NIH\_MGC\_256 Mus musculus cDNA clone  
IMAGE:30939911 5', mRNA sequence.  
ACCESSION COB15646  
VERSION COB15646.1 GI:51035046  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
TITLE NIH-MGC  
AUTHORS  
JOURNAL  
COMMENT  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
CDNA Library Preparation: Expressed Genomics  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: NDAM1201 row: c column: 24  
High quality sequence stop: 650.  
Location/Qualifiers  
1..739  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"

```

/clone="IMAGE:30939911"
/lab host="DH10B Tona"
/clone.lib="NIH_MGC_256"
/notes="Organ: oocyte; Vector: pExpress-1; Site_1: EcoRV;
Site_2: NotI; cDNA was primed using oligo-dT primer:
5'-pGACTGATGCTAGATCGGAGCGGCGCC(1)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary/library (normalized primary library is NIH_MGC_257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"

```

## ORIGIN

```

Query Match 7.9%; Score 320; DB 7; Length 739;
Best Local Similarity 69.5%; Pred. No. 1.3e-76;
Matches 448; Conservative 0; Mismatches 196; Indels 1; Gaps 1;

QY 2464 CTCAGTGAATGCTTGAAGAGTCTCCAGTGCCTGCGAGAACTGATCTGAGAGACTGT 2523
DB 1 CTTGGGAGTCCCTTGAAGTCAATGCTCTACCTGCAAAAGTTGATCTGAGCAACTGT 60
QY 2524 GGCATCAACGACGAGGCTTCCAGAGTGCCTCAAGCCCTGTCAGCAACCGAGCTTG 2583
DB 61 GGCCTCAACGACGAGGCTTCCAGAGTGCCTCAAGCCCTTTCAGCAACCGAGACTTG 120
QY 2584 ACACACCTGTGCTTCCAAACACAGCCCTGGGAGAGAGTGAATCTACTGTGTGCA 2643
DB 121 ACACACCTGTGCTTCCAAACACAGCCCTGGGAGAGAGTGAATCTACTGTGTGCA 180
QY 2644 TCCATGAGGCTTCCCACTGATGCTGAGAGGCTGATGATCAATGAGCCAGCTGAC 2703
DB 181 TTCTGAGAGATTCAGAAATGTCTCTCCAGCGGCTGATCTAATCACTGCAACATTGTA 240
QY 2704 ACGGCTGCTGTGTTTTCTTTCACCTTGGCTTATGAGTCACTGAGCTGACGACTG 2763
DB 241 GATGATGCTTATGCTCTCTGCGCAATGAGACTTGCACAAACAAAGCTGACCACTG 300
QY 2764 AGCCTTACGATTAACCTCTGTGAGAAAGATGCGGTGAAGCTTGTGCGAGTCAATG 2823
DB 301 AGCCTTACGATTAACCTCTGTGAGAAAGATGCGGTGAAGCTTGTGCGAGTCAATG 360
QY 2824 GAACCATCTTGTATCTCCAGAGCTGAGAGTGTAAAGTTCATCTCAACCGCGGTCG 2883
DB 321 GAACCATCTTGTATCTCCAGAGCTGAGAGTGTAAAGTTCATCTCAACCGCGGTCG 420
QY 2884 TGTGAGAGTGTCTCTGTGATCTGAGAGAGCAACCTGAGAGCTTGATCTCAAG 2943
DB 421 TGGGAGAGCTGCTGCTGATGATCAACCAACCAAGCACTTAAAGTTGATCTTGGT 480
QY 2944 GACATGCTTGTGAGAGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3003
DB 481 AACAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 3004 AGTGTCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3063
DB 541 AGTGTCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 3064 CTCTCTTGTGCTTCTTCTGCAACCGGATCTGACCACTGTAAC 3108
DB 601 TGTGATTTGGGCTATCTTTG-AACTCTGACCTGAACAGCTTAAC 644

```

## RESULT 15

```

BUE18831 642 bp mRNA linear EST 23-SEP-2002
LOCUS BUE18831/c
DEFINITION UI-H-FH1-BFL-1-19-0-UI-s1 NCI CGAP FH1 Homo sapiens cDNA clone
VERSION BUE18831.1 GI:23285046
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
1 (bases 1 to 642)
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

```

## FEATURES

```

source
1..642
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FH1-BFL-1-19-0-UI"
/tissue_type="Cell Line"
/dev_stage="Adult"
/lab host="DH10B (Life Technologies)"

```

```

/clone.lib="NCI CGAP_FH1"
/notes="Organ: Chondrosarcoma; Vector: pT7T3-Pac
(pharmacia) with a modified polylinker; Site_1: EcoR I;
Site_2: Not I; NCI CGAP FH1 is a normalized cDNA library
obtained from a cell line derived from grade I
chondrosarcoma tissue. The library was constructed and
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT7T3-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AGATCCGGC. The cell line was provided by Dr. James Martin
from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Cell Line C8 - Grade 1
Chondrosarcoma
TAG_LIB=UI-H-FH1
TAG_SEQ=AGATCCGGC"

```

## ORIGIN

```

Query Match 7.8%; Score 313.6; DB 5; Length 642;
Best Local Similarity 98.5%; Pred. No. 7.3e-75;
Matches 327; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 2981 GCGAGGAGCTGAAGAAAGAAAGAGTGTCTGAGAGAGCTGGGTTGAAGGATGTGAC 3040
DB 642 GCGAGGAGCTGAAGAAAGAAAGAGTGTCTGAGAGAGCTGGGTTGAAGGATGTGAC 584
QY 3041 TGACTTCTGATGTGTGAGGACACTCTCTGCGCCCTTCTTCTGCAACCGGATCTGACA 3100
DB 583 TGACTTCTGATGTGTGAGGACACTCTCTGCGCCCTTCTTCTGCAACCGGATCTGACA 524
QY 3101 GTCTTAACCTGTGCAAGTAATTAATCTCACTCCCAAGAGATGATGAAGTGTGTGGCT 3160
DB 523 GTCTTAACCTGTGCAAGTAATTAATCTCACTCCCAAGAGATGATGAAGTGTGTGGCT 464
QY 3161 TTGCTGTCCCAAGCTTAATTAAGATTAATTTGGGCTGTGAAATGAGGATACCTGTGC 3220
DB 463 TTGCTGTCCCAAGCTTAATTAAGATTAATTTGGGCTGTGAAATGAGGATACCTGTGC 404
QY 3221 AATTAAGAGAGCTGTGAGAGAGTGAAGTCACTCAAGCCCGAGTGTGTAATTTGACGTA 3280
DB 403 AATTAAGAGAGCTGTGAGAGAGTGAAGTCACTCAAGCCCGAGTGTGTAATTTGACGTA 344
QY 3281 GTTGGCATTTCTTTGATGAAGATGACGACAC 3312

```

Db 343 |||||  
TTGGCATTCTTTGATGAGATGACCGGTAC 312

Search completed: July 19, 2005, 01:37:18  
Job time : 12247 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 19, 2005, 12:03:24 ; Search time 32 Seconds

(without alignments)  
3135.261 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344  
Sequence: 1 MEGGXSLTFSSYGLQWCLYE.....DDHSGVMSLGAAGLEGIVS 1344Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size: 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	0.7	209	US-09-388-221B-19	Sequence 19, Appl
2	9	0.7	1399	US-09-388-221B-4	Sequence 4, Appl
3	9	0.7	1424	US-09-388-221B-12	Sequence 12, Appl
4	9	0.7	1443	US-09-388-221B-6	Sequence 6, Appl
5	9	0.7	1454	US-09-388-221B-10	Sequence 10, Appl
6	9	0.7	1473	US-09-388-221B-2	Sequence 2, Appl
7	8	0.6	150	US-09-270-767-32603	Sequence 32603, A
8	8	0.6	150	US-09-270-767-47820	Sequence 47820, A
9	8	0.6	296	US-09-253-991A-22293	Sequence 22293, A
10	8	0.6	321	US-09-902-540-10274	Sequence 10274, A
11	8	0.6	326	US-09-902-540-9957	Sequence 9957, Ap
12	8	0.6	388	US-09-248-796A-17824	Sequence 17824, A
13	8	0.6	661	US-09-371-338-7	Sequence 7, Appl
14	7	0.5	59	US-09-270-767-5972	Sequence 5972, A
15	7	0.5	60	US-09-543-681A-4574	Sequence 4574, Ap
16	7	0.5	60	US-09-513-999C-7527	Sequence 7527, Ap
17	7	0.5	63	US-09-513-999C-6481	Sequence 6481, Ap
18	7	0.5	94	US-09-270-767-59952	Sequence 59952, A
19	7	0.5	108	US-09-216-393B-4	Sequence 4, Appl
20	7	0.5	110	US-09-543-681A-6458	Sequence 6458, Ap
21	7	0.5	111	US-09-621-976-7083	Sequence 7083, Ap
22	7	0.5	117	US-10-101-464A-182	Sequence 182, App
23	7	0.5	118	US-09-270-767-41140	Sequence 41140, A
24	7	0.5	118	US-09-270-767-56356	Sequence 56356, A
25	7	0.5	126	US-09-640-211A-2108	Sequence 2108, App
26	7	0.5	128	US-09-733-210-647	Sequence 647, App
27	7	0.5	128	US-09-513-999C-7418	Sequence 7418, App

28	7	0.5	130	US-10-101-464A-780	Sequence 780, App
29	7	0.5	151	US-09-270-767-33903	Sequence 33903, A
30	7	0.5	174	US-09-270-767-44506	Sequence 44506, A
31	7	0.5	176	US-09-270-767-32581	Sequence 32581, A
32	7	0.5	176	US-09-270-767-47798	Sequence 47798, A
33	7	0.5	180	US-09-489-039A-9331	Sequence 9331, Ap
34	7	0.5	183	US-09-489-039A-7626	Sequence 7626, Ap
35	7	0.5	187	US-09-716-129-58	Sequence 58, Appl
36	7	0.5	189	US-08-779-764A-28	Sequence 28, Appl
37	7	0.5	189	US-08-779-764A-29	Sequence 29, Appl
38	7	0.5	189	US-08-779-764A-30	Sequence 30, Appl
39	7	0.5	189	US-09-563-456-28	Sequence 28, Appl
40	7	0.5	189	US-09-563-456-29	Sequence 29, Appl
41	7	0.5	189	US-09-563-456-30	Sequence 30, Appl
42	7	0.5	202	US-09-270-767-42656	Sequence 42656, A
43	7	0.5	220	US-09-902-540-12909	Sequence 12909, A
44	7	0.5	221	US-09-489-039A-12495	Sequence 12495, A
45	7	0.5	231	US-09-489-039A-8999	Sequence 8999, Ap

## ALIGNMENTS

```
RESULT 1
US-09-388-221B-19
Sequence 19, Application US/09388221B
Patent No. 6818750
GENERAL INFORMATION:
APPLICANT: Reed, John C.
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
FILE REFERENCE: P-LJ 3650
CURRENT FILING DATE: 1999-09-01
CURRENT APPLICATION NUMBER: US/09/388,221B
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 209
TYPE: PRT
ORGANISM: Homo sapiens
US-09-388-221B-19

Query Match          0.7%; Score 9; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      260 SRPERLLFI 268
Db      72 SRPERLLFI 80

RESULT 2
US-09-388-221B-4
Sequence 4, Application US/09388221B
Patent No. 6818750
GENERAL INFORMATION:
APPLICANT: Reed, John C.
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
FILE REFERENCE: P-LJ 3650
CURRENT FILING DATE: 1999-09-01
CURRENT APPLICATION NUMBER: US/09/388,221B
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1399
TYPE: PRT
ORGANISM: Homo sapiens
US-09-388-221B-4

Query Match          0.7%; Score 9; DB 4; Length 1399;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 400 SRPERLLFI 408

## RESULT 3

US-09-388-221B-12  
; Sequence 12, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-LJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; CURRENT FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 1424  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct  
US-09-388-221B-12

Query Match 0.7%; Score 9; DB 4; Length 1424;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

## RESULT 4

US-09-388-221B-6  
; Sequence 6, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-LJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; CURRENT FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1443  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-388-221B-6

Query Match 0.7%; Score 9; DB 4; Length 1443;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

## RESULT 5

US-09-388-221B-10  
; Sequence 10, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-LJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; CURRENT FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10

; LENGTH: 1454.

; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct  
US-09-388-221B-10

Query Match 0.7%; Score 9; DB 4; Length 1454;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

## RESULT 6

US-09-388-221B-2  
; Sequence 2, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-LJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; CURRENT FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1473  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-388-221B-2

Query Match 0.7%; Score 9; DB 4; Length 1473;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

## RESULT 7

US-09-270-767-32603  
; Sequence 32603, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 32603  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-32603

Query Match 0.6%; Score 8; DB 4; Length 150;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1197 AELEGLGS 1204  
Db 100 AELEGLGS 107

## RESULT 8

US-09-270-767-47820  
; Sequence 47820, Application US/09270767

```

; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 47820
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47820

Query Match          0.6%; Score 8; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1197 ALEGLGS 1204
Db      100 ALEGLGS 107

RESULT 9
US-09-252-991A-22293
; Sequence 22293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22293
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22293

Query Match          0.6%; Score 8; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      849 QSLASALV 856
Db      106 QSLASALV 113

RESULT 10
US-09-902-540-10274
; Sequence 10274, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10274
; LENGTH: 321
```

```

; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10274

Query Match          0.6%; Score 8; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      194 GKSIGKS 201
Db      158 GKSIGKS 165

RESULT 11
US-09-902-540-9957
; Sequence 9957, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9957
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-9957

Query Match          0.6%; Score 8; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      554 RRPLEVLL 561
Db      120 RRPLEVLL 127

RESULT 12
US-09-248-796A-17824
; Sequence 17824, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Ketch Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17824
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17824

Query Match          0.6%; Score 8; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      302 LIRKVLIP 309
Db      47 LIRKVLIP 54
```

```

RESULT 13
US-09-371-338-7
; Sequence 7, Application US/09371338
; Patent No. 6613959
; GENERAL INFORMATION:
; APPLICANT: Sheen, Jen
; APPLICANT: Kovun, Yelena V.
; APPLICANT: Chiu, Wan-ling
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING A MAPKKK
; FILE REFERENCE: 00786/366002
; CURRENT APPLICATION NUMBER: US/09/371,338
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/095,938
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-371-338-7

```

```

Query Match      0.6%; Score 8; DB 4; Length 661;
Best Local Similarity 100.0%; Pred. No. 99;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      274 DLGSVLNN 281
      |||||
Db      337 DLGSVLNN 344

```

```

RESULT 14
US-09-270-767-57972
; Sequence 57972, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57972
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57972

```

```

Query Match      0.5%; Score 7; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches      7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      475 NNILLPD 481
      |||||
Db      53 NNILLPD 59

```

```

RESULT 15
US-09-543-681A-4574
; Sequence 4574, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709 1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706

```

```

; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4574
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4574

```

```

Query Match      0.5%; Score 7; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches      7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      804 LKSLSLA 810
      |||||
Db      44 LKSLSLA 50

```

```

Search completed: July 19, 2005, 12:19:48
Job time : 35 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2005, 12:11:18 ; Search time 124 Seconds

(without alignments)  
4207.459 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344  
Sequence: 1 MEGKSLTFSSYGLQWCLYE.....DDHSGVMSLGAAGLEGLVS 1344

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1736639 seqs, 388188149 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1736639

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10E\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1344	US-10-124-498-6	Sequence 6, Appli
2	1344	100.0	1344	US-10-066-521-6	Sequence 6, Appli
3	1006	74.9	1162	US-10-216-645-2	Sequence 2, Appli
4	1006	74.9	1162	US-10-416-642-1	Sequence 1, Appli
5	988	73.5	1143	US-10-216-645-4	Sequence 4, Appli
6	968	72.0	1033	US-10-092-900A-348	Sequence 348, App
7	602	44.8	1200	US-10-399-443-24	Sequence 24, Appli
8	602	44.8	1200	US-10-677-943-24	Sequence 24, Appli
9	602	44.8	1200	US-10-860-761-4	Sequence 96, Appli
10	544	40.5	682	US-10-407-866-96	Sequence 96, Appli
11	158	11.8	385	US-10-399-443-2	Sequence 2, Appli

12	158	11.8	385	15	US-10-677-943-2	Sequence 2, Appli
13	149	11.1	258	15	US-10-399-443-4	Sequence 4, Appli
14	149	11.1	258	15	US-10-677-943-4	Sequence 4, Appli
15	12	0.9	190	10	US-09-965-621-63	Sequence 63, Appli
16	12	0.9	190	10	US-10-407-866-63	Sequence 63, Appli
17	12	0.9	190	16	US-10-781-294-63	Sequence 63, Appli
18	12	0.9	321	15	US-10-407-866-116	Sequence 116, App
19	12	0.9	449	14	US-10-028-374-2	Sequence 2, Appli
20	12	0.9	449	14	US-10-183-770-2	Sequence 2, Appli
21	12	0.9	635	15	US-10-407-866-90	Sequence 90, Appli
22	12	0.9	1004	15	US-10-108-260A-3161	Sequence 3161, Ap
23	12	0.9	1027	15	US-10-407-866-68	Sequence 68, Appli
24	12	0.9	1035	10	US-09-965-621-24	Sequence 24, Appli
25	12	0.9	1035	15	US-10-407-866-24	Sequence 24, Appli
26	12	0.9	1035	16	US-10-781-294-24	Sequence 18, Appli
27	12	0.9	1061	14	US-10-124-498-18	Sequence 18, Appli
28	12	0.9	1061	14	US-10-066-521-18	Sequence 5, Appli
29	12	0.9	1111	14	US-10-216-645-5	Sequence 6, Appli
30	12	0.9	1111	15	US-10-399-443-6	Sequence 6, Appli
31	12	0.9	1111	15	US-10-677-943-6	Sequence 29387, A
32	9	0.7	80	14	US-10-029-386-29387	Sequence 257211, A
33	9	0.7	87	16	US-10-425-115-257211	Sequence 34582, A
34	9	0.7	162	9	US-09-864-761-34582	Sequence 4432, Ap
35	9	0.7	192	10	US-09-764-891-4432	Sequence 13890, A
36	9	0.7	224	17	US-10-732-923-13890	Sequence 23227, A
37	9	0.7	284	15	US-10-424-599-223227	Sequence 114, App
38	9	0.7	303	14	US-10-306-762-114	Sequence 92, Appli
39	9	0.7	764	15	US-10-407-866-92	Sequence 4, Appli
40	9	0.7	1399	9	US-09-388-221-4	Sequence 12, Appli
41	9	0.7	1424	9	US-09-388-221-12	Sequence 2, Appli
42	9	0.7	1429	9	US-09-996-617-2	Sequence 2, Appli
43	9	0.7	1429	9	US-09-931-071-2	Sequence 3, Appli
44	9	0.7	1429	14	US-10-029-347-3	Sequence 26, Appli
45	9	0.7	1429	14	US-10-029-347-26	

## ALIGNMENTS

RESULT 1  
US-10-124-498-6  
Sequence 6, Application US/10124498  
Publication No. US2003001983A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Wang, Weiye  
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
FILE REFERENCE: 07334-367001  
CURRENT APPLICATION NUMBER: US/10/124,498  
PRIOR FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: 10/066,521  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: 60/318,645  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/265,231  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 1344  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-124-498-6  
Query Match 100.0%; Score 1344; DB 14; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 MEGKSLTFSSYGLQWCLYEIDKEEFTFKELAKKSSSTTCSIPQFEIENAVECTAL 60  
Db 1 MEGKSLTFSSYGLQWCLYEIDKEEFTFKELAKKSSSTTCSIPQFEIENAVECTAL 60

QY 61 LHHEYYGASLAWATSISIFENNNLRTLSEKARDMKKISQAMEOGATAETEOEISOA 120  
DB 61 LHHEYYGASLAWATSISIFENNNLRTLSEKARDMKKISQAMEOGATAETEOEISOA 120  
QY 121 MEQEGATAETEOEGHGDITWDYKSHVMTKFAEEDVRSPFENTAADMEQOTLAGAFDS 180  
DB 121 MEQEGATAETEOEGHGDITWDYKSHVMTKFAEEDVRSPFENTAADMEQOTLAGAFDS 180  
QY 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
DB 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
QY 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLANNDYTKLCKDMAEKOPFTLIR 300  
DB 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLANNDYTKLCKDMAEKOPFTLIR 300  
QY 301 SLRKRVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISGEORIHLLERIGEHQKT 360  
DB 301 SLRKRVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISGEORIHLLERIGEHQKT 360  
QY 361 QGLRAIMNNRELLDQCQVPVAGSLICVALQLODVVGESVAPPNQTLTGLHAAFFVHQLTP 420  
DB 361 QGLRAIMNNRELLDQCQVPVAGSLICVALQLODVVGESVAPPNQTLTGLHAAFFVHQLTP 420  
QY 421 RGVVRCINLEBERVYLKFFCRMAVEGVNNRKSVPFGDDLAVOGLGESELRALFHMNILLP 480  
DB 421 RGVVRCINLEBERVYLKFFCRMAVEGVNNRKSVPFGDDLAVOGLGESELRALFHMNILLP 480  
QY 481 DSHCEYVTFPHLSLQDFCAALYYVLEGLIEPALCPLYVEKTKSMELKQAGFIHSLW 540  
DB 481 DSHCEYVTFPHLSLQDFCAALYYVLEGLIEPALCPLYVEKTKSMELKQAGFIHSLW 540  
QY 541 MKRFLFGVSEBVRPLEVLLGCPVPLGVKOKLHMVSLLGQOPATTPGTLAEHCLF 600  
DB 541 MKRFLFGVSEBVRPLEVLLGCPVPLGVKOKLHMVSLLGQOPATTPGTLAEHCLF 600  
QY 601 ETQDEFPYRLALNSFOEYWLPIQNLIDLIASFCLOHPYLRKIRVDYKGIFFRDESABA 660  
DB 601 ETQDEFPYRLALNSFOEYWLPIQNLIDLIASFCLOHPYLRKIRVDYKGIFFRDESABA 660  
QY 661 CPVVPDLMBRDKTLIEQWEDFCSMLGTHPHLRQDLGSSILITERAKTLCALRHPTCKI 720  
DB 661 CPVVPDLMBRDKTLIEQWEDFCSMLGTHPHLRQDLGSSILITERAKTLCALRHPTCKI 720  
QY 721 QTLMBRNOITPGVOHLMRIWANNNRSLNMGTHLKEEDVRMACEALKHPKCLBSLR 780  
DB 721 QTLMBRNOITPGVOHLMRIWANNNRSLNMGTHLKEEDVRMACEALKHPKCLBSLR 780  
QY 781 LDCCGLTHACYLKIISQILTTSPSLKSLAGNKVTDQVMP,SDALRVSOQALQTLIED 840  
DB 781 LDCCGLTHACYLKIISQILTTSPSLKSLAGNKVTDQVMP,SDALRVSOQALQTLIED 840  
QY 841 CGITATGCGSALASALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHSCISQRLMNOCHL 900  
DB 841 CGITATGCGSALASALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHSCISQRLMNOCHL 900  
QY 901 DTAGGGLTALALMGSM,THLSLSNVPEDNGVLLCEWMBRPSCHLQDLBYKCHLTA 960  
DB 901 DTAGGGLTALALMGSM,THLSLSNVPEDNGVLLCEWMBRPSCHLQDLBYKCHLTA 960  
QY 961 CCESTLSCVTSRSHKSLDLTDNALGDGVAALCEGLKOKNSVLTFLGLKAKG,LTSDCE 1020  
DB 961 CCESTLSCVTSRSHKSLDLTDNALGDGVAALCEGLKOKNSVLTFLGLKAKG,LTSDCE 1020  
QY 1021 ALSTLALSCNRHLTSLNIVQNNFSPKGMK,CSAFACPTSN,IIIGLWMOY,PPVOIRKLE 1080  
DB 1021 ALSTLALSCNRHLTSLNIVQNNFSPKGMK,CSAFACPTSN,IIIGLWMOY,PPVOIRKLE 1080  
QY 1081 EVOLLKPRVVIDGSHWSFDEDDRHKIGLTFRLPESRAMP,CALLMKNMBQKRVSLAGD 1140  
DB 1081 EVOLLKPRVVIDGSHWSFDEDDRHKIGLTFRLPESRAMP,CALLMKNMBQKRVSLAGD 1140

QY 1141 FKSSTRPAKSLCLATANGESQVDNVEOSSPOPAAGTEHKKDKMLSVYSGAMSETAELE 1200  
DB 1141 FKSSTRPAKSLCLATANGESQVDNVEOSSPOPAAGTEHKKDKMLSVYSGAMSETAELE 1200  
QY 1201 GLGSNSADHDHGGWAMSLGRLSRLG,CPYLTMTTAVCPGHMERL,SGRGCLNSADHSG 1260  
DB 1201 GLGSNSADHDHGGWAMSLGRLSRLG,CPYLTMTTAVCPGHMERL,SGRGCLNSADHSG 1260  
QY 1261 VWSLGAAGLEGGLVNSNADHDSGVAMSLGAAGLEGGLVNSNADHDSGVAMSLGAAGLEGGLV 1320  
DB 1261 VWSLGAAGLEGGLVNSNADHDSGVAMSLGAAGLEGGLVNSNADHDSGVAMSLGAAGLEGGLV 1320  
QY 1321 SNSADHDSGVAMSLGAAGLEGGLVNSNADHDSGVAMSLGAAGLEGGLV 1344  
DB 1321 SNSADHDSGVAMSLGAAGLEGGLVNSNADHDSGVAMSLGAAGLEGGLV 1344

RESULT 2  
US-10-066-521-6  
; Sequence 6, Application US/1006521  
; Publication No. US20030027757A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; APPLICANT: Wang, Weiye  
; APPLICANT: Blatcher, Maria  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-334001  
; CURRENT APPLICATION NUMBER: US/10/066, 521  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 60/318, 645  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/265, 231  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-066-521-6

Query Match 100.0%; Score 1344; DB 14; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGDKSLTFSSYGIQWCLYELDKKEFOTFKELKKKSSSESTTCSIPOFEIENNAVECLAL 60  
DB 1 MEGDKSLTFSSYGIQWCLYELDKKEFOTFKELKKKSSSESTTCSIPOFEIENNAVECLAL 60  
QY 61 LHHEYYGASLAWATSISIFENNNLRTLSEKARDMKKISQAMEOGATAETEOEISOA 120  
DB 61 LHHEYYGASLAWATSISIFENNNLRTLSEKARDMKKISQAMEOGATAETEOEISOA 120  
QY 121 MEQEGATAETEOEGHGDITWDYKSHVMTKFAEEDVRSPFENTAADMEQOTLAGAFDS 180  
DB 121 MEQEGATAETEOEGHGDITWDYKSHVMTKFAEEDVRSPFENTAADMEQOTLAGAFDS 180  
QY 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
DB 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
QY 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLANNDYTKLCKDMAEKOPFTLIR 300  
DB 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLANNDYTKLCKDMAEKOPFTLIR 300  
QY 301 SLRKRVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISGEORIHLLERIGEHQKT 360  
DB 301 SLRKRVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISGEORIHLLERIGEHQKT 360  
QY 361 QGLRAIMNNRELLDQCQVPVAGSLICVALQLODVVGESVAPPNQTLTGLHAAFFVHQLTP 420  
DB 361 QGLRAIMNNRELLDQCQVPVAGSLICVALQLODVVGESVAPPNQTLTGLHAAFFVHQLTP 420



```

QY 421 RGVNRCLNTERVYLKPRCRMAVEGVNRKSVPGDDLMNGLGSESLRALFHHNILLP 480
DB 421 RGVNRCLNTERVYLKPRCRMAVEGVNRKSVPGDDLMNGLGSESLRALFHHNILLP 480
QY 481 DSHCEYYTFPHLSLQDFCALYYVLGELIEPALCPLVEKTKSMELKQAFHISLM 540
DB 481 DSHCEYYTFPHLSLQDFCALYYVLGELIEPALCPLVEKTKSMELKQAFHISLM 540
QY 541 MKRPLFGLVSEVDYRPLEVLLGCPVPLGVKQKLLHNVSLLGQOPNATTPGDTLDAFHCLE 600
DB 541 MKRPLFGLVSEVDYRPLEVLLGCPVPLGVKQKLLHNVSLLGQOPNATTPGDTLDAFHCLE 600
QY 601 ETQDEPFRALANSQOEWMPLINOMLDIASSFCLOHCYLAIRKIVDVYKGIIPRDESAA 660
DB 601 ETQDEPFRALANSQOEWMPLINOMLDIASSFCLOHCYLAIRKIVDVYKGIIPRDESAA 660
QY 661 CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDSSILTERAMKTLCAKLRHPTCKI 720
DB 661 CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDSSILTERAMKTLCAKLRHPTCKI 720
QY 721 QTLMRNAQITPVOHLMRIYANANRLSLNLGTHLKEEDVRMACEALKRHKCLLESIR 780
DB 721 QTLMRNAQITPVOHLMRIYANANRLSLNLGTHLKEEDVRMACEALKRHKCLLESIR 780
QY 781 LDCCGITHACYLIKIOILTTSPSLKSLAGKNTVDQVMPISDALRVSOCLQKLLIED 840
DB 781 LDCCGITHACYLIKIOILTTSPSLKSLAGKNTVDQVMPISDALRVSOCLQKLLIED 840
QY 841 CGITANGCQSLASALVSNRSLTHLCLSNNSLGNBGNVLCRSMRLPHGSLQRLMNQCHL 900
DB 841 CGITANGCQSLASALVSNRSLTHLCLSNNSLGNBGNVLCRSMRLPHGSLQRLMNQCHL 900
QY 901 DTAGCGFLALALMGNMWLTHLSLSPMPVEDNGVKLLCEVWRPESCHLQDLIELVKCHLTA 960
DB 901 DTAGCGFLALALMGNMWLTHLSLSPMPVEDNGVKLLCEVWRPESCHLQDLIELVKCHLTA 960
QY 961 CCEBSICVTSRSRHLKSLDLTDNALGDGVVALCEGLKOKNSVLTRLGKACGLTSDCE 1020
DB 961 CCEBSICVTSRSRHLKSLDLTDNALGDGVVALCEGLKOKNSVLTRLGKACGLTSDCE 1020
QY 1021 ALSLASCNRHLTSLNLYONNFSPKGMKLCGAPCPTSNLQITGLMKQYVQIRKYLE 1080
DB 1021 ALSLASCNRHLTSLNLYONNFSPKGMKLCGAPCPTSNLQITGLMKQYVQIRKYLE 1080
QY 1081 EYQOLKPRVVIDSGMHSFEDDRHKIGLTFRLPESRAMPCALLMGNNPEQKRVSLLAGD 1140
DB 1081 EYQOLKPRVVIDSGMHSFEDDRHKIGLTFRLPESRAMPCALLMGNNPEQKRVSLLAGD 1140
QY 1141 FKSSSTRFPAKSLCLATANGESQKRVNVBOSSPQMAGTEHKODKMLSVGSGAMSETAIE 1200
DB 1141 FKSSSTRFPAKSLCLATANGESQKRVNVBOSSPQMAGTEHKODKMLSVGSGAMSETAIE 1200
QY 1201 GIGNSADHDGGMAMSLGRELSRGLCPTVMTTAAYVCGHMERLGSRCWCLNSADDSG 1260
DB 1201 GIGNSADHDGGMAMSLGRELSRGLCPTVMTTAAYVCGHMERLGSRCWCLNSADDSG 1260
QY 1261 VSMISGAAGLEGLVSNADDSGVAMSLGAAGLEGLVSNADDSGVAMSLGAAGLEGLV 1320
DB 1261 VSMISGAAGLEGLVSNADDSGVAMSLGAAGLEGLVSNADDSGVAMSLGAAGLEGLV 1320
QY 1321 SNSADDSGVAMSLGAAGLEGLV 1344
DB 1321 SNSADDSGVAMSLGAAGLEGLV 1344

```

```

RESULT 3
US-10-216-645-2
; Sequence 2, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LEBST, MONIKA

```

```

; APPLICANT: PETERS-KOTTIG, MICHAEL
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN MATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216, 645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-645-2

Query Match      74.9%; Score 1006; DB 14; Length 1162;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1006; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 ISQAMEQCATPAETEEQEOISQAMEQCATPAETEEQCHGDYWDYKSHWTKFAEEDV 157
DB 152 ISQAMEQCATPAETEEQEOISQAMEQCATPAETEEQCHGDYWDYKSHWTKFAEEDV 211
QY 158 RRSFENTADMPENQTLGAFDSDRMGFRPTVVLHGSGIGKSALARIVLCMAQGGLY 217
DB 212 RRSFENTADMPENQTLGAFDSDRMGFRPTVVLHGSGIGKSALARIVLCMAQGGLY 271
QY 218 QGMFSYVFFLVREMQKKESSVTEFISREMPDSQAPTEIIMSBERILLIIDGFDIDGS 277
DB 272 QGMFSYVFFLVREMQKKESSVTEFISREMPDSQAPTEIIMSBERILLIIDGFDIDGS 331
QY 278 VLANDTKLCKDMAEKOPFTLIRSLRKVLLPESFLIVTVADVTEKJKSEVSPRYLLV 337
DB 332 VLANDTKLCKDMAEKOPFTLIRSLRKVLLPESFLIVTVADVTEKJKSEVSPRYLLV 391
QY 338 RGISGEQRIHLLERIGEGHOKTOGIRAIMNRRELDCCVPAYGSLCVALLQODVVE 397
DB 392 RGISGEQRIHLLERIGEGHOKTOGIRAIMNRRELDCCVPAYGSLCVALLQODVVE 451
QY 398 SVAPENQTLTGILAAFFVHQLTPRGVNRCLNTERVYLKPRCRMAVEGVNRKSVFPGD 457
DB 452 SVAPENQTLTGILAAFFVHQLTPRGVNRCLNTERVYLKPRCRMAVEGVNRKSVFPGD 511
QY 458 DLMVQGLSESLRALFHHNILLPDSHCEYYTFPHLSLQDFCALYYVLGELIEPALCP 517
DB 512 DLMVQGLSESLRALFHHNILLPDSHCEYYTFPHLSLQDFCALYYVLGELIEPALCP 571
QY 518 LYVEKTKSMELKQAFHISLMKRRFLGLVSEVDYRPLEVLLGCPVPLGVKQKLLHNV 577
DB 572 LYVEKTKSMELKQAFHISLMKRRFLGLVSEVDYRPLEVLLGCPVPLGVKQKLLHNV 631
QY 578 SLLGQOPNATTPGDTLDAFHCLEFETQDEPFRALANSQOEWMPLINOMLDIASSFCLOH 637
DB 632 SLLGQOPNATTPGDTLDAFHCLEFETQDEPFRALANSQOEWMPLINOMLDIASSFCLOH 691
QY 638 CPYLKIRKIVDVKGIFPRDESAAACPVVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDGL 697
DB 692 CPYLKIRKIVDVKGIFPRDESAAACPVVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDGL 751
QY 698 SSILTERAMKTLCAKLRHPTCKIOTLMFRNAQITPVOHLMRIYANANRLSLNLGTHL 757
DB 752 SSILTERAMKTLCAKLRHPTCKIOTLMFRNAQITPVOHLMRIYANANRLSLNLGTHL 811
QY 758 KEBDVRMACEALKRHKCLLESIRLDCCGITHACYLIKIOILTTSPSLKSLAGKNTVDQ 817
DB 812 KEBDVRMACEALKRHKCLLESIRLDCCGITHACYLIKIOILTTSPSLKSLAGKNTVDQ 871
QY 818 GVMPLSDALRVSOCLQKLLIEDCGITATGQSLASALVSNRSLTHLCLSNNSLGNBGN 877
DB 872 GVMPLSDALRVSOCLQKLLIEDCGITATGQSLASALVSNRSLTHLCLSNNSLGNBGN 931
QY 878 LLCRSMRLPHGSLQRLMNQCHLDTAGCGFLALALMGNMWLTHLSLSPMPVEDNGVKLLC 937

```

Db 932 LLCRSMRLPHCSLQGLMLNQCGLDTAGCGFLALALMGNSWLTHLSLSNMPVEDNGVLLC 991  
Qy 938 EVMREPSCHLODELVYKCHLTAACCESSLSVCVSRBRHLKSLDTLTMALGDGVALLCGL 997  
Db 992 EVMREPSCHLODELVYKCHLTAACCESSLSVCVSRBRHLKSLDTLTMALGDGVALLCGL 1051  
Qy 998 KOKNSVLTRLGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1057  
Db 1052 KOKNSVLTRLGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1111  
Qy 1058 TSNLIQILGMKQYVQVQIRKLLBEVQLLKPRVVIDGSMHSFDEDDR 1103  
Db 1112 TSNLIQILGMKQYVQVQIRKLLBEVQLLKPRVVIDGSMHSFDEDDR 1157

RESULT 4  
US-10-416-642-1  
; Sequence 1, Application US/10416642  
; Publication No. US20040043452A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: RAKUMAR, Jayajakmi  
; APPLICANT: ARVIZU, Chandra  
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0842 PCT  
; CURRENT APPLICATION NUMBER: US/10/416,642  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/249,407  
; PRIOR FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 1162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CD1  
US-10-416-642-1

Query Match 74.9%; Score 1006; DB 15; Length 1162;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1006; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 98 ISQAMBOGATAETAEBOEISQAMBOGATAETEOGHGDTWYKSHWMTKFAEEDV 157  
Db 152 ISQAMBOGATAETAEBOEISQAMBOGATAETEOGHGDTWYKSHWMTKFAEEDV 211  
Qy 158 RRSFNTAADWPEMOTLAGAFSDRMGFRPTVVLHGKSGIGKSALARIYLCMAQGGLY 217  
Db 212 RRSFNTAADWPEMOTLAGAFSDRMGFRPTVVLHGKSGIGKSALARIYLCMAQGGLY 271  
Qy 218 QGMFSYVFFLPVREKORKESSVTFEISREWPDQAPTEIWSRPERLLFTIIDGFDIGS 277  
Db 272 QGMFSYVFFLPVREKORKESSVTFEISREWPDQAPTEIWSRPERLLFTIIDGFDIGS 331  
Qy 278 VLNNDTKLCCKMAEKOPPTLLRSLLRKVLLPESFLIYTVVDVGEKLSKSEVSPRYLLV 337  
Db 332 VLNNDTKLCCKMAEKOPPTLLRSLLRKVLLPESFLIYTVRVGTEKLSKSEVSPRYLLV 391  
Qy 338 RGISGEORIHLLERIGIEGHOKTOGLRAIMNNRELLDCCQVPAVGSLLCVALQLODVGE 397  
Db 392 RGISGEORIHLLERIGIEGHOKTOGLRAIMNNRELLDCCQVPAVGSLLCVALQLODVGE 451  
Qy 398 SVAPFNQTLTGHAAPVTHQLTGRGVVRCCLNLEERVYLKFCRMAVEGWNRRKSVFDGD 457  
Db 452 SVAPFNQTLTGHAAPVTHQLTGRGVVRCCLNLEERVYLKFCRMAVEGWNRRKSVFDGD 511  
Qy 458 DLMVGLESSELRALFHNHLLPDSHCERYTFFHLSTODPCAALYYTLLEGITRPAACP 517  
Db 512 DLMVGLESSELRALFHNHLLPDSHCERYTFFHLSTODPCAALYYTLLEGITRPAACP 571

Qy 518 LYVEKTRSMELKQAGFHISHLMKREFLFGIVSEVDVRRPLEVLLGCPVPLGVKOKLHMV 577  
Db 572 LYVEKTRSMELKQAGFHISHLMKREFLFGIVSEVDVRRPLEVLLGCPVPLGVKOKLHMV 631  
Qy 578 SLGQOPNATTPGDTLDAFHCLEFETOKEFVRLALNSFOEVMPLINONLDLIASSFCLQH 637  
Db 632 SLGQOPNATTPGDTLDAFHCLEFETOKEFVRLALNSFOEVMPLINONLDLIASSFCLQH 691  
Qy 638 CPVLRKTRVDYKGIFFRDESAAECVYVPLMWRDRTLLEBOWEDCSMLGTHPHLRQDLG 697  
Db 692 CPVLRKTRVDYKGIFFRDESAAECVYVPLMWRDRTLLEBOWEDCSMLGTHPHLRQDLG 751  
Qy 698 SSILTERAMKTLCAKLHPCTCKIOTLMFRNAQITPGVQHLMRIYMANRNLRSNLIGSTHL 757  
Db 752 SSILTERAMKTLCAKLHPCTCKIOTLMFRNAQITPGVQHLMRIYMANRNLRSNLIGSTHL 811  
Qy 758 KEDVIRMACEALKRPKCLLESRLDCCGLTHACTYKTSOILITSPSKSLAGNKYTDQ 817  
Db 812 KEDVIRMACEALKRPKCLLESRLDCCGLTHACTYKTSOILITSPSKSLAGNKYTDQ 871  
Qy 818 GWMPLSDALRYSCALQKLLIEDCGITATGCGQSALASLVNSRSLTHLCLSNNSLGNBYN 877  
Db 872 GWMPLSDALRYSCALQKLLIEDCGITATGCGQSALASLVNSRSLTHLCLSNNSLGNBYN 931  
Qy 878 LLCRSMRLPHCSLQGLMLNQCGLDTAGCGFLALALMGNSWLTHLSLSNMPVEDNGVLLC 937  
Db 932 LLCRSMRLPHCSLQGLMLNQCGLDTAGCGFLALALMGNSWLTHLSLSNMPVEDNGVLLC 991  
Qy 938 EVMREPSCHLODELVYKCHLTAACCESSLSVCVSRBRHLKSLDTLTMALGDGVALLCGL 997  
Db 992 EVMREPSCHLODELVYKCHLTAACCESSLSVCVSRBRHLKSLDTLTMALGDGVALLCGL 1051  
Qy 998 KOKNSVLTRLGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1057  
Db 1052 KOKNSVLTRLGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1111  
Qy 1058 TSNLIQILGMKQYVQVQIRKLLBEVQLLKPRVVIDGSMHSFDEDDR 1103  
Db 1112 TSNLIQILGMKQYVQVQIRKLLBEVQLLKPRVVIDGSMHSFDEDDR 1157

RESULT 5  
US-10-216-645-4  
; Sequence 4, Application US/10216645  
; Publication No. US20030125282A1  
; GENERAL INFORMATION:  
; APPLICANT: WEISS, BERTRAM  
; APPLICANT: LESSL, MONIKA  
; APPLICANT: PETERS-KOTTIG, MICHAEL  
; APPLICANT: BECKMANN, GEORG  
; TITLE OF INVENTION: HUMAN MATER PROTEINS  
; FILE REFERENCE: SCH-1910  
; CURRENT APPLICATION NUMBER: US/10/216,645  
; CURRENT FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-216-645-4

Query Match 73.5%; Score 988; DB 14; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 988; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 116 EISQAMBOGATAETAEBOEISQAMBOGATAETEOGHGDTWYKSHWMTKFAEEDVRRSFNTAADWPEMOTLA 175  
Db 151 EISQAMBOGATAETAEBOEISQAMBOGATAETEOGHGDTWYKSHWMTKFAEEDVRRSFNTAADWPEMOTLA 210  
Qy 176 GAFSDRMGFRPTVVLHGKSGIGKSALARIYLCMAQGGLYQGMFSYVFFLPVREKOR 235

211 GAFSDRNGFRRTVVLHGKSGIGKSLARRIVLCMAOGGVLQGMFSYVFLPVREMRK 270  
226 KESSTTERISREWPDSQAPVTEINSPERLFIITGFPDLSGVANNPTCLCDMAEKOP 295  
271 KESSTTERISREWPDSQAPVTEINSPERLFIITGFPDLSGVANNPTCLCDMAEKOP 330  
296 FTLLSLRKVLPSPFLVTVRDVTEKSKSEWSPRLVIRGISEGRIHLLEERGIG 355  
331 FTLLSLRKVLPSPFLVTVRDVTEKSKSEWSPRLVIRGISEGRIHLLEERGIG 390  
356 EHOKTQGLRAIMNNELLDDQCVPAVGSILCVALQLODVGSVAPENOTLTLHAFAVF 415  
391 EHOKTQGLRAIMNNELLDDQCVPAVGSILCVALQLODVGSVAPENOTLTLHAFAVF 450  
416 HOLTGRGVRRCLNLEERVVTLKRCMAVEGWNRKSVFDGDDLMVQGLSESELALPHM 475  
451 HOLTGRGVRRCLNLEERVVTLKRCMAVEGWNRKSVFDGDDLMVQGLSESELALPHM 510  
476 NILPDSHCEEYTFPHSLDPCALVYVLEGLTEIPALCPLVYEKTRSMELKQAGRH 535  
511 NILPDSHCEEYTFPHSLDPCALVYVLEGLTEIPALCPLVYEKTRSMELKQAGRH 570  
536 IHSLMKRFPLGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLGQOPNATTPDITDA 595  
571 IHSLMKRFPLGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLGQOPNATTPDITDA 630  
596 FHCLFETDKEFVRILNLSFOEVRPLINONDLIASFCLQHPYLKIRIVDKGIFPRD 655  
631 FHCLFETDKEFVRILNLSFOEVRPLINONDLIASFCLQHPYLKIRIVDKGIFPRD 690  
656 ESASECPVPLMNRKTLIEQWEDPCMLGTHPHLRQULGSSILTERAMTKLCAKLRH 715  
691 ESASECPVPLMNRKTLIEQWEDPCMLGTHPHLRQULGSSILTERAMTKLCAKLRH 750  
716 PTCKIQTLMFRNAQITPGVQHLRIVMANRNLRSNLGTHLKEEDVMAACEALKHPKCL 775  
751 PTCKIQTLMFRNAQITPGVQHLRIVMANRNLRSNLGTHLKEEDVMAACEALKHPKCL 810  
776 LESRLDCCGTLTHACYLKISQILTTSPSLKSLAGNVTGCVPLSDALRVSCALQK 835  
811 LESRLDCCGTLTHACYLKISQILTTSPSLKSLAGNVTGCVPLSDALRVSCALQK 870  
836 LILBECGTTANGCCSLAALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHCSLQRLML 895  
871 LILBECGTTANGCCSLAALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHCSLQRLML 930  
896 NOCHUDTAGCGFLAALMGNSWLTHLSLNMNVEDNGVYLCEVNRBPSCHLQDELEVYC 955  
931 NOCHUDTAGCGFLAALMGNSWLTHLSLNMNVEDNGVYLCEVNRBPSCHLQDELEVYC 990  
956 HLTAAACSLSCVLSRSRHLKSLDLTDNALGDGVAALCEGLKQKNSVLTSLGLKACGLT 1015  
991 HLTAAACSLSCVLSRSRHLKSLDLTDNALGDGVAALCEGLKQKNSVLTSLGLKACGLT 1050  
1016 SDCCGALSLAISCNRHLTSLNVQNNFSPKGMKCSAPACTSMLQITIGLKMOPYVQI 1075  
1051 SDCCGALSLAISCNRHLTSLNVQNNFSPKGMKCSAPACTSMLQITIGLKMOPYVQI 1110  
1076 RKLLEEVOLLKPRVYIDGSMHSFDEDDR 1103  
1111 RKLLEEVOLLKPRVYIDGSMHSFDEDDR 1138

APPLICANT: Pena, Carol B.A.  
APPLICANT: Li, Li  
APPLICANT: Zernusen, Bryan D.  
APPLICANT: Gueev, Vladimir Y.  
APPLICANT: Ji, Weizhen  
APPLICANT: Gorman, Linda  
APPLICANT: Miller, Charles E.  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Paturajan, Meera  
APPLICANT: Gangolli, Baha A.  
APPLICANT: Verneet, Corine A.M.  
APPLICANT: Guo, Xiaojia Sasha  
APPLICANT: Tchernev, Velizar T.  
APPLICANT: Fernandes, Elma R.  
APPLICANT: Casman, Stacie J.  
APPLICANT: Malyankar, Uriel M.  
APPLICANT: Gerlach, Valerie  
APPLICANT: Liu, Yi  
APPLICANT: Anderson, David W.  
APPLICANT: Spaderna, Steven K.  
APPLICANT: Caterton, Elina  
APPLICANT: Leite, Mario W.  
APPLICANT: Zhong, Haihong  
APPLICANT: Alsobrook, John P.  
APPLICANT: Lepley, Denise M.  
APPLICANT: Rieger, Daniel K.  
APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: NO. US2004004382A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C  
CURRENT APPLICATION NUMBER: US/10/092,900A  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: USN 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USN 60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: USN 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: USN 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USN 60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USN 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: USN 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: USN 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: USN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USN 60/287,424  
Remaining Prior Application data removed - See file wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 348  
LENGTH: 1033  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-900A-348

Query Match 72.0%; Score 968; DB 15; Length 1033;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 968; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

136 HGGDTWDYKSHVMTKFAEEDVRRSPENTADWPEMOTLAGAFSDRNGFRRTVVLHGK 195  
61 HGGDTWDYKSHVMTKFAEEDVRRSPENTADWPEMOTLAGAFSDRNGFRRTVVLHGK 120  
196 SGIGKSLARRIVLCMAOGGVLQGMFSYVFLPVREMRKESSTTERISREWPDSQAPV 255  
121 SGIGKSLARRIVLCMAOGGVLQGMFSYVFLPVREMRKESSTTERISREWPDSQAPV 180  
256 TEINSPERLFIITGFPDLSGVANNPTCLCDMAEKOPPTLLRSLLRKVLPESEFLIV 315

181 TEIMSRPERLLFTIDGFDLGSVLNNDTKLCKDMAEKOPFTLLIRLSLRKVLPPESFLIV 240  
QY 316 TWRDGTSEKLSSEVSPRYLLVRGISEGQRHLLERIGEHOKTQGRALMNNRELLDQ 375  
Db 241 TWRDGTSEKLSSEVSPRYLLVRGISEGQRHLLERIGEHOKTQGRALMNNRELLDQ 300  
QY 376 CQVPVAGSLICVALQLODVGVESVAPFNQTLTGLHAAVFHQLTGRGVRRCLNLEERV 435  
Db 301 CQVPVAGSLICVALQLODVGVESVAPFNQTLTGLHAAVFHQLTGRGVRRCLNLEERV 360  
QY 436 LKRFRCMAVEGVWNRKSVFDDGLMWQGLGESELRALFHMNTLLPDSHCEBYTTFHLSL 495  
Db 361 LKRFRCMAVEGVWNRKSVFDDGLMWQGLGESELRALFHMNTLLPDSHCEBYTTFHLSL 420  
QY 496 ODFCALYYVLEGLIEPALCPLYVEKTKRSMELKQAGPHIHSIMMKKFLGLVSEDDR 555  
Db 421 ODFCALYYVLEGLIEPALCPLYVEKTKRSMELKQAGPHIHSIMMKKFLGLVSEDDR 480  
QY 556 PLEVLLGCPVPLGAVKQKLLHWVSLIGQOPNATTPGDTLDAFHCLFETQDKFVRLLANSF 615  
Db 481 PLEVLLGCPVPLGAVKQKLLHWVSLIGQOPNATTPGDTLDAFHCLFETQDKFVRLLANSF 540  
QY 616 QEWPLPINQNLDLIASSFCLOHCPYLKRIKIVDKGIFPRDESABACPVVPLMRDKTLIE 675  
Db 541 QEWPLPINQNLDLIASSFCLOHCPYLKRIKIVDKGIFPRDESABACPVVPLMRDKTLIE 600  
QY 676 BOWEPFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPCTCKIOTLMFRNAQITTPVQ 735  
Db 601 BOWEPFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPCTCKIOTLMFRNAQITTPVQ 660  
QY 736 HLMRIWMANRNLRSINTLGGTHLKEEDVMAACEALKHPCKLLESRLDDCCGLTHACYLKIS 795  
Db 661 HLMRIWMANRNLRSINTLGGTHLKEEDVMAACEALKHPCKLLESRLDDCCGLTHACYLKIS 720  
QY 796 QILTTSPSIKSISLAGNKVTTDQGVNPLSDALRVSCALOKLILEDCGTTATGCCSLASAL 855  
Db 721 QILTTSPSIKSISLAGNKVTTDQGVNPLSDALRVSCALOKLILEDCGTTATGCCSLASAL 780  
QY 856 VSNRELTLCLSNNSLGNBGNVLLCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGN 915  
Db 781 VSNRELTLCLSNNSLGNBGNVLLCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGN 840  
QY 916 SWLTLHLSLSPNVEEDNGVYLLCEVWRBPSCHLQDLIELVKCHLTAACCSLSGVISRSRL 975  
Db 841 SWLTLHLSLSPNVEEDNGVYLLCEVWRBPSCHLQDLIELVKCHLTAACCSLSGVISRSRL 900  
QY 976 KSLDLITDVALGQVYALCEGLKQKNSVLTIRGLKACGLTSDCCALSLASCNHRTSL 1035  
Db 901 KSLDLITDVALGQVYALCEGLKQKNSVLTIRGLKACGLTSDCCALSLASCNHRTSL 960  
QY 1036 NLVONNPSFKGMKLCSAFACPTSNLQIIGLMKMOYPOAIRKLEEVYLLKPRVVIDGSW 1095  
Db 961 NLVONNPSFKGMKLCSAFACPTSNLQIIGLMKMOYPOAIRKLEEVYLLKPRVVIDGSW 1020  
QY 1096 HSFDEDDR 1103  
Db 1021 HSFDEDDR 1028

RESULT 7  
US-10-399-443-24  
; Sequence 24, Application US/10399443  
; Publication No. US20040028669A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785

; CURRENT APPLICATION NUMBER: US/10/399, 443  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/241, 510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 1200  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-399-443-24

Query Match 44.8%; Score 602; DB 15; Length 1200;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 98 ISQAMEQGAATAETEEQOISQAMEQGAATAETEEQGHGDDTWYKSHWTKFAEEDV 157  
Db 190 ISQAMEQGAATAETEEQOISQAMEQGAATAETEEQGHGDDTWYKSHWTKFAEEDV 249  
QY 158 RRSFENTAADWPEMOTIAGAFDSRWGFRPTTVVHLKSGSIGKSALARIVLCWAQGLY 217  
Db 250 RRSFENTAADWPEMOTIAGAFDSRWGFRPTTVVHLKSGSIGKSALARIVLCWAQGLY 309  
QY 218 QGMFSYFPLPYRMOKKSSVTEFTSRMPDSQAVTEIMSRPERLLFTIDGFDLGS 277  
Db 310 QGMFSYFPLPYRMOKKSSVTEFTSRMPDSQAVTEIMSRPERLLFTIDGFDLGS 369  
QY 278 VLANDTKLCKDMAEKOPFTLLIRLSLRKVLPPESFLIVTVADVTEKLSKEVSPRYLLV 337  
Db 370 VLANDTKLCKDMAEKOPFTLLIRLSLRKVLPPESFLIVTVADVTEKLSKEVSPRYLLV 429  
QY 338 RGISEGQRHLLERIGEHOKTQGRALMNNRELLDQCVPAVAGSLICVALQLODVVGE 397  
Db 430 RGISEGQRHLLERIGEHOKTQGRALMNNRELLDQCVPAVAGSLICVALQLODVVGE 489  
QY 398 SVAFPNQTLTGLHAAVFHQLTGRGVRRCLNLEERVVLRFCGMAVEGVWNRKSVDDG 457  
Db 490 SVAFPNQTLTGLHAAVFHQLTGRGVRRCLNLEERVVLRFCGMAVEGVWNRKSVDDG 549  
QY 458 DLMQGLGESELRALFHMNTLLPDSHCEBYTTFHLSLQDFCALYYVLEGLIEPALCP 517  
Db 550 DLMQGLGESELRALFHMNTLLPDSHCEBYTTFHLSLQDFCALYYVLEGLIEPALCP 609  
QY 518 LYVEKTKRSMELKQAGPHIHSIMMKRFLGLVSEDDRPLEVLLGCPVPLGAVKQKLLHW 577  
Db 610 LYVEKTKRSMELKQAGPHIHSIMMKRFLGLVSEDDRPLEVLLGCPVPLGAVKQKLLHW 669  
QY 578 SLIGQOPNATTPGDTLDAFHCLFETQDKFVRLLANSFOEWMPLPINQNLDLIASSFCLOH 637  
Db 670 SLIGQOPNATTPGDTLDAFHCLFETQDKFVRLLANSFOEWMPLPINQNLDLIASSFCLOH 729  
QY 638 CPYLKRIKIVDKGIFPRDESABACPVVPLMRDKTLIEBQWEDFCMLGTHPHLRQDLG 697  
Db 730 CPYLKRIKIVDKGIFPRDESABACPVVPLMRDKTLIEBQWEDFCMLGTHPHLRQDLG 789  
QY 698 SSILTERAMKTLCAKLRHPCTCKIOTLMFRNAQITTPVQHLMRIWMANRNLRSINTLGGTHL 757  
Db 790 SSILTERAMKTLCAKLRHPCTCKIOTLMFRNAQITTPVQHLMRIWMANRNLRSINTLGGTHL 849  
QY 758 KEEVNRMACEALKHPCKLLESRLDDCCGLTHACVLYKSQILTTSPSIKSISLAGNKVTTDQ 817  
Db 850 KEEVNRMACEALKHPCKLLESRLDDCCGLTHACVLYKSQILTTSPSIKSISLAGNKVTTDQ 909  
QY 818 GWPFLDALRVSCALOKLILEDCGTTATGCCSLASALVNSRSLTLCLSNNSLGNBGNV 877  
Db 910 GWPFLDALRVSCALOKLILEDCGTTATGCCSLASALVNSRSLTLCLSNNSLGNBGNV 969  
QY 878 LKCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGNSWLTLHLSLSPNVEEDNGVYLLC 937  
Db 970 LKCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGNSWLTLHLSLSPNVEEDNGVYLLC 1029

QY 938 EVMREPSCHLADLELVKCHLTAACCESLSCVSRSHRLKSLDITDNLGSGVALACBGL 997  
DB 1030 EVMREPSCHLADLELVKCHLTAACCESLSCVSRSHRLKSLDITDNLGSGVALACBGL 1089  
QY 998 KOKNSVLTRELGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCGSAFACP 1057  
DB 1090 KOKNSVLTRELGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCGSAFACP 1149  
QY 1058 TSNLOITIGLWKQYVQVQIRKLEEVQOLKPRVVDGSHWSPDEDDR 1103  
DB 1150 TSNLOITIGLWKQYVQVQIRKLEEVQOLKPRVVDGSHWSPDEDDR 1195

RESULT 8  
US-10-677-943-24  
; Sequence 24, Application US/10677943  
; Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zhi-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT APPLICATION NUMBER: US/10/677,943  
; CURRENT FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US02/09776  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 24  
; LENGTH: 1200  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-677-943-24

Query Match 44.8%; Score 602; DB 15; Length 1200;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 98 ISQAMEOGATTAETEEBOEISQAMEOGATTAETEEBOGSDTYMDYKSHWTKFAEBEDV 157  
DB 190 ISQAMEOGATTAETEEBOEISQAMEOGATTAETEEBOGSDTYMDYKSHWTKFAEBEDV 249  
QY 158 RRSFNTAADWPEMOTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARIVLCWAQGGLY 217  
DB 250 RRSFNTAADWPEMOTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARIVLCWAQGGLY 309  
QY 218 QGMESYVFFLPVREMOKKESVTEFISREWPSQAPVTEIMSRERLLFIIDGDDGLS 277  
DB 310 QGMESYVFFLPVREMOKKESVTEFISREWPSQAPVTEIMSRERLLFIIDGDDGLS 369  
QY 278 VLNDTKLCKDMAEKQPPFTLIRSLIRKVLPESEFLITVRVDVGEKLSKSEVSPRYLLV 337  
DB 370 VLNDTKLCKDMAEKQPPFTLIRSLIRKVLPESEFLITVRVDVGEKLSKSEVSPRYLLV 429  
QY 338 RGISGEORIHLLERIGEHOHTQGLRAIMNRELDDCCVAVASLICVALQLDDVGE 397  
DB 430 RGISGEORIHLLERIGEHOHTQGLRAIMNRELDDCCVAVASLICVALQLDDVGE 489  
QY 398 SVAPNQTLLTGLHAFVHQUTPRGVVRCNLLEERVVLKRCRAVAGVWNRKSVFPGD 457  
DB 490 SVAPNQTLLTGLHAFVHQUTPRGVVRCNLLEERVVLKRCRAVAGVWNRKSVFPGD 549  
QY 458 DLMVGLGESELRALFHNHLLPDSHCCEYTPFPHLSLDDFCALYYVLGELIPALCP 517  
DB 550 DLMVGLGESELRALFHNHLLPDSHCCEYTPFPHLSLDDFCALYYVLGELIPALCP 609

QY 518 LYVEKTRSMELKQAGFIHSHLMKRFLFGLVSEDEVRRPLEVILGCPVPLGKQOLHMY 577  
DB 610 LYVEKTRSMELKQAGFIHSHLMKRFLFGLVSEDEVRRPLEVILGCPVPLGKQOLHMY 669  
QY 578 SLGQOPNATTPGDTLDAPHCLFETQDEFYRLALNSFOEVLPIPNOMLIDLIASFCLQH 637  
DB 670 SLGQOPNATTPGDTLDAPHCLFETQDEFYRLALNSFOEVLPIPNOMLIDLIASFCLQH 729  
QY 638 CPLYLRKIRVDYKGIFFRDESAEACPVPLMRDKTLIEQWEDFCSMGTGPHRLQLDLG 697  
DB 730 CPLYLRKIRVDYKGIFFRDESAEACPVPLMRDKTLIEQWEDFCSMGTGPHRLQLDLG 789  
QY 698 SSILTERAMKTLCAALRHPTCKIOTLMFRNAQITPGVQLRWIRVANNRLSLNGTTL 757  
DB 790 SSILTERAMKTLCAALRHPTCKIOTLMFRNAQITPGVQLRWIRVANNRLSLNGTTL 849  
QY 758 KEEDVRACEMALKHPKCLLESRLDCCGLTACVYKISQILTTSPSLKSLIAGNKVTDQ 817  
DB 850 KEEDVRACEMALKHPKCLLESRLDCCGLTACVYKISQILTTSPSLKSLIAGNKVTDQ 909  
QY 818 GVMPLSDALRVSOQALQTLIEDCGITATGCOISALAVNSRLTTLCLSNNSLGNBGN 877  
DB 910 GVMPLSDALRVSOQALQTLIEDCGITATGCOISALAVNSRLTTLCLSNNSLGNBGN 969  
QY 878 LLCRSMLRPHCSLQRLMNOCHDITAGGFLALALMGWSLTHLSLNNPVEDNGVLLC 937  
DB 970 LLCRSMLRPHCSLQRLMNOCHDITAGGFLALALMGWSLTHLSLNNPVEDNGVLLC 1029  
QY 938 EVMREPSCHLADLELVKCHLTAACCESLSCVSRSHRLKSLDITDNLGSGVALACBGL 997  
DB 1030 EVMREPSCHLADLELVKCHLTAACCESLSCVSRSHRLKSLDITDNLGSGVALACBGL 1089  
QY 998 KOKNSVLTRELGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCGSAFACP 1057  
DB 1090 KOKNSVLTRELGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCGSAFACP 1149  
QY 1058 TSNLOITIGLWKQYVQVQIRKLEEVQOLKPRVVDGSHWSPDEDDR 1103  
DB 1150 TSNLOITIGLWKQYVQVQIRKLEEVQOLKPRVVDGSHWSPDEDDR 1195

RESULT 9  
US-10-860-761-4  
; Sequence 4, Application US/10860761  
; Publication No. US2004024875A1  
; GENERAL INFORMATION:  
; APPLICANT: WYETH  
; TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS  
; FILE REFERENCE: AM101318  
; CURRENT APPLICATION NUMBER: US/10/860,761  
; CURRENT FILING DATE: 2004-06-03  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 1200  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-860-761-4

Query Match 44.8%; Score 602; DB 16; Length 1200;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 98 ISQAMEOGATTAETEEBOEISQAMEOGATTAETEEBOGSDTYMDYKSHWTKFAEBEDV 157  
DB 190 ISQAMEOGATTAETEEBOEISQAMEOGATTAETEEBOGSDTYMDYKSHWTKFAEBEDV 249  
QY 158 RRSFNTAADWPEMOTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARIVLCWAQGGLY 217  
DB 250 RRSFNTAADWPEMOTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARIVLCWAQGGLY 309  
QY 218 QGMESYVFFLPVREMOKKESVTEFISREWPSQAPVTEIMSRERLLFIIDGDDGLS 277

```

Db 310 QGMFSYVFLPVREMQRKSSVTEFISREWPDSQAPVTEIMSRPERLLFTIDGFDLGS 369
Qy 278 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVTEKLSKSEVSPRYLLV 337
Db 370 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVTEKLSKSEVSPRYLLV 429
Qy 338 RGISGEORIHLLERIGIEHOKTQGLRAIMNNRELLDCCOVPAYSLICVALQLODVGE 397
Db 430 RGISGEORIHLLERIGIEHOKTQGLRAIMNNRELLDCCOVPAYSLICVALQLODVGE 489
Qy 398 SVAPNQTLLTGLHAFFHOLTFRGVVRCINLEERVVLRKFCMAVAGVNNRKSVPDGD 457
Db 490 SVAPNQTLLTGLHAFFHOLTFRGVVRCINLEERVVLRKFCMAVAGVNNRKSVPDGD 549
Qy 458 DLMVQGLSESELRALFHMNIILLPDSCHEEYTFPHLSLODFCALYYVLEGLTEPALCP 517
Db 550 DLMVQGLSESELRALFHMNIILLPDSCHEEYTFPHLSLODFCALYYVLEGLTEPALCP 609
Qy 518 LYVEKTKSMELKQAGFIHISLMKRFLGLVSEDEVRRPLEVLGCPVPLGVKQKLHMV 577
Db 610 LYVEKTKSMELKQAGFIHISLMKRFLGLVSEDEVRRPLEVLGCPVPLGVKQKLHMV 669
Qy 578 SLGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYWLPINONLIDLIASSFCLOH 637
Db 670 SLGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYWLPINONLIDLIASSFCLOH 729
Qy 638 CPYLKTIKIVDKGIFRDESAACGVPLMMRDKTLIEOWEDFCSMGTGPHLRQLDG 697
Db 730 CPYLKTIKIVDKGIFRDESAACGVPLMMRDKTLIEOWEDFCSMGTGPHLRQLDG 789
Qy 698 SSILTERAMKTLICARHPCKIQTLMFRNAQITPGVOHLMRIWANNELSLNIGTHTL 757
Db 790 SSILTERAMKTLICARHPCKIQTLMFRNAQITPGVOHLMRIWANNELSLNIGTHTL 849
Qy 758 KEEVDYMACEALKHPCILLESIRLDCCGITHACYIKISQIILTTSPSLKSLIAGKVTQD 817
Db 850 KEEVDYMACEALKHPCILLESIRLDCCGITHACYIKISQIILTTSPSLKSLIAGKVTQD 909
Qy 818 GMPPLSDALRVSGCALOKLILIEDGITTATGCGSLASALVSNRSLTHLCLSNNSLGENEYV 877
Db 910 GMPPLSDALRVSGCALOKLILIEDGITTATGCGSLASALVSNRSLTHLCLSNNSLGENEYV 969
Qy 878 LLCRSNRLPHCSLQRLMLNOCHLDTAGCGFLALALMGNSWLTHLSLSNMPVEDNGVKLLC 937
Db 970 LLCRSNRLPHCSLQRLMLNOCHLDTAGCGFLALALMGNSWLTHLSLSNMPVEDNGVKLLC 1029
Qy 938 EYMRPESCHLODELVKCHLTAAACESLSCVTSRSHLKSJDLITNALGDGVAAALCEGL 997
Db 1030 EYMRPESCHLODELVKCHLTAAACESLSCVTSRSHLKSJDLITNALGDGVAAALCEGL 1089
Qy 998 KOKNSVLRRLGKAGAGLSDCEALSLALSCNRHLTSINLVONNSPKGMMLCSAFACP 1057
Db 1090 KOKNSVLRRLGKAGAGLSDCEALSLALSCNRHLTSINLVONNSPKGMMLCSAFACP 1149
Qy 1058 TSNLIQIIGLMKQYFVQIRKLLEEVOLKPRVVIDGSMHSPEDDR 1103
Db 1150 TSNLIQIIGLMKQYFVQIRKLLEEVOLKPRVVIDGSMHSPEDDR 1195

RESULT 10
US-10-407-866-96
; Sequence 96, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAD Domain-Containing Polypeptides,
; FILE REFERENCE: 66554-10 (Ij 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
```

```

; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 682
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-96

Query Match 40.5%; Score 544; DB 15; Length 682;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 ISQAMEGATPAETEOEISQAMEGATPAETEOGHGGDTWDYKSHWTKFAEBDV 157
Db 139 ISQAMEGATPAETEOEISQAMEGATPAETEOGHGGDTWDYKSHWTKFAEBDV 198
Qy 158 RRSFENTADMPENQTTAGAFSDRWGRPRPTVTHGSGIGKSALARRIYLCMAOGGLY 217
Db 199 RRSFENTADMPENQTTAGAFSDRWGRPRPTVTHGSGIGKSALARRIYLCMAOGGLY 258
Qy 218 QGMFSYVFLPVREMQRKSSVTEFISREWPDSQAPVTEIMSRPERLLFTIDGFDLGS 277
Db 259 QGMFSYVFLPVREMQRKSSVTEFISREWPDSQAPVTEIMSRPERLLFTIDGFDLGS 318
Qy 278 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVTEKLSKSEVSPRYLLV 337
Db 319 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVTEKLSKSEVSPRYLLV 378
Qy 338 RGISGEORIHLLERIGIEHOKTQGLRAIMNNRELLDCCOVPAYSLICVALQLODVGE 397
Db 379 RGISGEORIHLLERIGIEHOKTQGLRAIMNNRELLDCCOVPAYSLICVALQLODVGE 438
Qy 398 SVAPNQTLLTGLHAFFHOLTFRGVVRCINLEERVVLRKFCMAVAGVNNRKSVPDGD 457
Db 439 SVAPNQTLLTGLHAFFHOLTFRGVVRCINLEERVVLRKFCMAVAGVNNRKSVPDGD 498
Qy 458 DLMVQGLSESELRALFHMNIILLPDSCHEEYTFPHLSLODFCALYYVLEGLTEPALCP 517
Db 499 DLMVQGLSESELRALFHMNIILLPDSCHEEYTFPHLSLODFCALYYVLEGLTEPALCP 558
Qy 518 LYVEKTKSMELKQAGFIHISLMKRFLGLVSEDEVRRPLEVLGCPVPLGVKQKLHMV 577
Db 559 LYVEKTKSMELKQAGFIHISLMKRFLGLVSEDEVRRPLEVLGCPVPLGVKQKLHMV 618
Qy 578 SLGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYWLPINONLIDLIASSFCLOH 637
Db 619 SLGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYWLPINONLIDLIASSFCLOH 678
Qy 638 CPYL 641
Db 679 CPYL 682

RESULT 11
US-10-399-443-2
; Sequence 2, Application US/10399443
; Publication No. US20040028693A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
```



;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 2  
;; LENGTH: 385  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-399-443-2

Query Match 11.8%; Score 158; DB 15; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.1e-142;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 AEKOPFTLIRSLRKVLPESEFLIVTVADVGTETKLSVSPRYLVKSGEORIHLL 349  
DB 39 AEKOPFTLIRSLRKVLPESEFLIVTVADVGTETKLSVSPRYLVKSGEORIHLL 98  
QY 350 LERGIQEHQKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 409  
DB 99 LERGIQEHQKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 158  
QY 410 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCEMAVEGV 447  
DB 159 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCEMAVEGV 196

## RESULT 12

US-10-677-943-2  
; Sequence 2, Application US/10677943  
; Publication No. US20040072297A1

;; GENERAL INFORMATION:  
;; APPLICANT: The Government of the United States of America as  
;; APPLICANT: represented by the Secretary of the Department of Health and  
;; APPLICANT: Human Services  
;; APPLICANT: Nelson, Lawrence  
;; APPLICANT: Tong, Zhi-Bin  
;; TITLE OF INVENTION: Human Gene Critical to Fertility  
;; FILE REFERENCE: 4239-64790  
;; CURRENT APPLICATION NUMBER: US/10/677,943  
;; CURRENT FILING DATE: 2003-10-01  
;; PRIOR APPLICATION NUMBER: 60/241,510  
;; PRIOR FILING DATE: 2000-10-18  
;; PRIOR APPLICATION NUMBER: PCT/US02/09776  
;; PRIOR FILING DATE: 2002-03-29  
;; PRIOR APPLICATION NUMBER: PCT/US01/10981  
;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 42  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 2  
;; LENGTH: 385  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-677-943-2

Query Match 11.8%; Score 158; DB 15; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.1e-142;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 AEKOPFTLIRSLRKVLPESEFLIVTVADVGTETKLSVSPRYLVKSGEORIHLL 349  
DB 39 AEKOPFTLIRSLRKVLPESEFLIVTVADVGTETKLSVSPRYLVKSGEORIHLL 98  
QY 350 LERGIQEHQKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 409  
DB 99 LERGIQEHQKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 158  
QY 410 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCEMAVEGV 447  
DB 159 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCEMAVEGV 196

## RESULT 13

US-10-399-443-4

;; Sequence 4, Application US/10399443  
;; Publication No. US20040028669A1

;; GENERAL INFORMATION:  
;; APPLICANT: The Government of the United States of America, as Represented by the  
;; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
;; APPLICANT: Health  
;; APPLICANT: Nelson, Lawrence M.  
;; APPLICANT: Tong, Zhi-Bin  
;; APPLICANT: Nelson, Lawrence  
;; APPLICANT: Zhi-Bin, Tong  
;; TITLE OF INVENTION: Human Gene Critical to Fertility  
;; FILE REFERENCE: 4239-64785  
;; CURRENT APPLICATION NUMBER: US/10/399,443  
;; CURRENT FILING DATE: 2003-04-16  
;; PRIOR APPLICATION NUMBER: 60/241,510  
;; PRIOR FILING DATE: 2000-10-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/10981  
;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 4  
;; LENGTH: 258  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-399-443-4

Query Match 11.1%; Score 149; DB 15; Length 258;  
Best Local Similarity 100.0%; Pred. No. 3.6e-134;  
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 LTHLSLNNPYEDNKKVLCCEVMPRESCHDLEIVKCHLTPACCESISCVISRSRHKS 977  
DB 68 LTHLSLNNPYEDNKKVLCCEVMPRESCHDLEIVKCHLTPACCESISCVISRSRHKS 127  
QY 978 LDTLDMALGDGVVALCEGLKQKXSVLTRLGLKACGLTSDCEALSLALSCNRHLSINTL 1037  
DB 128 LDTLDMALGDGVVALCEGLKQKXSVLTRLGLKACGLTSDCEALSLALSCNRHLSINTL 187  
QY 1038 VQNNFSPKGMKLCGAFACPTSNLQITGL 1066  
DB 188 VQNNFSPKGMKLCGAFACPTSNLQITGL 216

## RESULT 14

US-10-677-943-4  
; Sequence 4, Application US/10677943  
; Publication No. US20040072297A1

;; GENERAL INFORMATION:  
;; APPLICANT: The Government of the United States of America as  
;; APPLICANT: represented by the Secretary of the Department of Health and  
;; APPLICANT: Human Services  
;; APPLICANT: Nelson, Lawrence  
;; APPLICANT: Tong, Zhi-Bin  
;; TITLE OF INVENTION: Human Gene Critical to Fertility  
;; FILE REFERENCE: 4239-64790  
;; CURRENT APPLICATION NUMBER: US/10/677,943  
;; CURRENT FILING DATE: 2003-10-01  
;; PRIOR APPLICATION NUMBER: 60/241,510  
;; PRIOR FILING DATE: 2000-10-18  
;; PRIOR APPLICATION NUMBER: PCT/US02/09776  
;; PRIOR FILING DATE: 2002-03-29  
;; PRIOR APPLICATION NUMBER: PCT/US01/10981  
;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 42  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 4  
;; LENGTH: 258  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-677-943-4

Query Match 11.1%; Score 149; DB 15; Length 258;  
Best Local Similarity 100.0%; Pred. No. 3.6e-134;

Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 LTHLSNPNVEDNGVKLLCEVMRPSCHLQDLRLVKCHLTRACCESLSCVTSRSHLKS 977  
|||||  
DB 68 LTHLSNPNVEDNGVKLLCEVMRPSCHLQDLRLVKCHLTRACCESLSCVTSRSHLKS 127  
|||||

QY 978 LDLTGNALGDGVALCEGLKQKNSVLTRLGLKACGLTSDCCALSLALSCNRHLSLNL 1037  
|||||  
DB 128 LDLTGNALGDGVALCEGLKQKNSVLTRLGLKACGLTSDCCALSLALSCNRHLSLNL 187  
|||||

QY 1038 VQNNFSPKGMKLCSAFACPTSNLQITGL 1066  
|||||  
DB 188 VQNNFSPKGMKLCSAFACPTSNLQITGL 216  
|||||

## RESULT 15

US-09-965-621-63  
; Sequence 63, Application US/09965621  
; Publication No. US20030077699A1

## GENERAL INFORMATION:

; APPLICANT: Reed, John C.  
; APPLICANT: Godzik, Adam  
; APPLICANT: Chu, Zhi-Liang  
; APPLICANT: Pawlowski, Krzysztof  
; APPLICANT: Fiorentino, Ioredana  
; APPLICANT: Ariza, Maria Eugenia  
; APPLICANT: Stehlik, Christian  
; TITLE OF INVENTION: PAD Domain-Containing Polypeptides,  
; FILE REFERENCE: P-LJ 4816  
; CURRENT APPLICATION NUMBER: US/09/965,621  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 09/671,760  
; PRIOR FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-965-621-63

Query Match 0.9%; Score 12; DB 10; Length 190;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 PERLFTIDGFD 273  
|||||  
DB 77 PERLFTIDGFD 88  
|||||

Search completed: July 19, 2005, 12:27:33  
Job time : 127 secs